Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **420,** 1825–1847 (2012) doi:10.1111/j.1365-2966.2011.19805.x

A combined optical and X-ray study of unobscured type 1 active galactic nuclei – I. Optical spectra and spectral energy distribution modelling

Chichuan Jin,[1⋆] Martin Ward,[1] Chris Done[1] and Jonathan Gelbord[1,2]

[1]*Department of Physics, University of Durham, South Road, Durham DH1 3LE*
[2]*Department of Astronomy & Astrophysics, Penn State University, PA 16801, USA*

Accepted 2011 September 12. Received 2011 September 12; in original form 2010 November 8

ABSTRACT

We present modelling and interpretation of the continuum and emission lines for a sample of 51 unobscured type 1 active galactic nuclei (AGNs). All of these AGNs have high-quality spectra from both *XMM–Newton* and the Sloan Digital Sky Survey. We extend the wavelength coverage where possible by adding simultaneous ultraviolet data from the OM onboard *XMM–Newton*. Our sample is selected based on low reddening in the optical and low gas columns implied by their X-ray spectra, except for one case, the broad absorption line quasar PG 1004+130. They also lack clear signatures for the presence of a warm absorber. Therefore, the observed characteristics of this sample are likely to be directly related to the intrinsic properties of the central engine.

To determine the intrinsic optical continuum, we subtract the Balmer continuum and all major emission lines (including Fe II). We also consider possible effects of contamination from the host galaxy. The resulting continuum is then used to derive the properties of the underlying accretion disc. We constrain the black hole masses from spectral fits of the Balmer emission lines and determine the best-fitting value from the modelling of broad-band spectral energy distributions (SEDs). In addition to the disc component, many of these SEDs also exhibit a strong soft X-ray excess, plus a power law extending to higher X-ray energies. We fit these SEDs by applying a new broad-band SED model which comprises accretion disc emission, low-temperature optically-thick Comptonization and a hard X-ray tail by introducing the concept of a corona radius. We find that in order to fit the data, the model often requires an additional long-wavelength optical continuum component, whose origin is discussed in this paper. We also find that the photorecombination edge of the Balmer continuum shifts and broadens beyond the standard limit of 3646 Å, implying an electron number density which is far higher than that in the broad-line-region clouds.

Our results indicate that the narrow-line type 1 Seyfert galaxies in this sample tend to have lower black hole masses, higher Eddington ratios, softer 2–10 keV band spectra, lower 2–10 keV luminosities and higher α_{ox}, compared with typical broad-line type 1 Seyfert galaxies, although their bolometric luminosities are similar. We illustrate these differences in properties by forming an average SED for three subsamples, based on the full width at half-maximum velocity width of the Hβ emission line.

Key words: accretion, accretion discs – galaxies: active.

1 INTRODUCTION

The spectral energy distribution (SED) of active galactic nuclei (AGNs) has been modelled for several decades. Initial studies focused on the infrared, optical and ultraviolet (UV) continua (e.g. Wills, Netzer & Wills 1985; Canalizo & Stockton 2001; Lacy et al. 2007). With the inclusion of X-ray data, it was possible to define the continuum on both sides of the UV/X-ray gap (imposed by Galactic photoelectric absorption), and so constrain the properties of the accretion disc (e.g. Ward et al. 1987; Elvis et al. 1994). Refinements to modelling the optical/UV continuum include subtraction of the complex blended features arising from permitted iron emission, the so-called small blue bump from the Balmer continuum, and

⋆E-mail: chichuan.jin@durham.ac.uk

contamination across the entire spectrum from a stellar component (Maoz et al. 1993; Boisson et al. 2000).

The observed spectral differences between various types of AGNs are not only due to selective absorption and orientation effects, as implied by the simplest version of AGN unification model (Antonucci 1993), but also result from a wide range in basic physical parameters, such as black hole mass and accretion rate (e.g. Boroson & Green 1992; Boller, Brandt & Fink 1996; Done & Gierliński 2005; Zhou et al. 2006). To better understand the accretion processes occurring close to the supermassive black hole, we construct broad-band SEDs. Galactic dust reddening, together with the intrinsic reddening of the AGN itself, attenuates the optical/UV band emission. Furthermore, photoelectric absorption from gas modifies the lower energy X-ray continuum. However, these factors can be quantified and corrected. Thereby, we can recover the intrinsic SED, except for the unobservable far-UV region. If we have reliable data on both sides of the energy gap between the UV and soft X-ray, we can apply a multicomponent model which spans across it.

1.1 Previous work

Many multiwavelength studies have been carried out previously. Puchnarewicz et al. (1992) studied the optical properties of 53 AGNs in Córdova et al.'s (1992) sample with ultrasoft X-ray excesses, and found that they tend to have narrower permitted lines than optically selected samples. Supporting this finding, Boller et al. (1996) studied *ROSAT*-selected AGNs with extremely soft X-ray spectra, and found that they tend to be narrow-line type 1 Seyfert galaxies (NLS1s). Correspondingly, they found that optically selected NLS1s often have large soft X-ray excesses. Walter & Fink (1993) combined soft X-ray and optical data for 58 type 1 Seyfert galaxies, and showed that their broad-band SEDs have a bump from the UV to soft X-rays, which is now referred to as the big blue bump (BBB). Grupe et al. (1998) and Grupe et al. (1999) used a sample of 76 bright, soft X-ray selected Seyfert galaxies with infrared data, optical spectra and soft X-ray spectra. Their results reinforced the connection between the optical and soft X-ray spectra, and confirmed the existence of strong BBB emission in these objects. Elvis et al. (1994) studied 47 quasars in a UV–soft X-ray sample, and derived the mean SEDs for radio-loud and radio-quiet sources. Recently, more detailed spectral models have been applied to broad-band SEDs including simultaneous optical/UV and X-ray observations which avoid potential problems caused by variability. Vasudevan & Fabian (2007) (hereinafter VF07) combined a disc and a broken power-law model to fit optical, far-UV and X-ray data for 54 AGNs. They found a well-defined relationship between the hard X-ray bolometric correction and the Eddington ratio. Brocksopp et al. (2006) analysed the data from *XMM–Newton*'s simultaneous EPIC (X-ray) and OM (optical/UV) observations for 22 Palomar Green (PG) quasars. Another sample consisting of 21 NLS1s and 13 broad-line AGNs was also defined using simultaneous data from *XMM–Newton*'s EPIC and OM monitor (Crummy et al. 2006). The SEDs of this sample were then fitted using various broad-band SED models such as a disc plus power-law model, disc reflection model and disc wind absorption model (Middleton, Done & Gierliński 2007). Vasudevan & Fabian (2009) derived SEDs using *XMM–Newton*'s simultaneous X-ray and optical/UV observations for 29 AGNs selected from Peterson et al.'s (2004) reverberation mapped sample. The well-constrained black hole masses available for this sample enabled them to fit a better constrained accretion disc model, combined with a power law, to the source's broad-band SEDs. Hence, they derived more reliable Eddington ratios.

1.2 Our AGN sample

In this paper, we define an X-ray/optically selected sample of 51 AGNs, all of which have low reddening (so excluding type 2 Seyfert galaxies and type 1.9/1.8 Seyfert galaxies), to construct SEDs ranging from about 0.9 μm to 10 keV. We also apply corrections for the permitted iron features, the Balmer continuum and stellar contribution, in order to model the non-stellar continuum free from emission-line effects. Included in this sample are a number of NLS1s, a subclass of AGNs whose permitted linewidths are comparable to those of forbidden lines. Their [O III] λ5007/Hβ ratio is also lower than the typical value of broad-line type 1 Seyfert galaxies (BLS1s) (Shuder & Osterbrock 1981; Osterbrock & Pogge 1985). For consistency with previous work, we classify AGNs in our sample as NLS1s if they have ratios of [O III] λ5007/Hβ < 3 and full width at half-maximum $\mathrm{FWHM_{H\beta}} < 2000\ \mathrm{km\,s^{-1}}$ (Goodrich 1989). We identify 10–12 NLS1s in our sample.[1]

All objects in our sample have high-quality optical spectra taken from the Sloan Digital Sky Survey (SDSS) DR7 (seventh data release of the SDSS), X-ray spectra from the *XMM–Newton* EPIC cameras, and in some cases simultaneous optical/UV photometric data points from the *XMM–Newton* OM monitor. Combining these data reduces the impact of intrinsic variability and provides a good estimate of the spectral shape in the optical, near-UV and X-ray regions. In addition, by analysing the SDSS spectra, we can derive the parameters of the principal optical emission lines and underlying continuum. An important result from reverberation mapping study is the correlation between black hole mass, monochromatic luminosity at 5100 Å and Hβ FWHM (e.g. Kaspi et al. 2000; Woo & Urry 2002; Peterson et al. 2004). We measure these quantities from the SDSS spectra, and then estimate black hole masses using this correlation.

Compared with previous work, a significant improvement of our study is that we employ a new broad-band SED model which combines disc emission, Comptonization and a high-energy power-law component in the context of an energetically self-consistent model for the accretion disc emission (Done et al. 2011, also see Section 5.2). By fitting this model to our data, we can reproduce the whole broad-band SED from the optical to X-ray. From this detailed SED fitting, we derive a number of interesting AGN properties such as the bolometric luminosity, Eddington ratio, hard X-ray slope and hard X-ray bolometric correction. Combining all the broad-band SED parameters with the optical parameters, we can provide further evidence for many previously suggested correlations, including all the correlations between the optical and X-ray claimed in previous work, plus many others such as the Hβ FWHM versus X-ray slope, black hole mass versus Eddington ratio, Fe II luminosity versus [O III] λ5007 emission-line luminosity and high excitation lines (e.g. [Fe VII] λ6087, [Fe X] λ6374) versus their ionizing flux (e.g. Boroson & Green 1992; Boller et al. 1996; Grupe et al. 1998; Grupe et al. 1999; Sulentic et al. 2000; Mullaney et al. 2009).

This paper is organized as follows. Section 2 describes the sample selection and data analysis procedures. The detailed spectral-fitting methods and results including Balmer line fitting, optical spectral fitting and broad-band SED fitting are each discussed in Sections 3, 4 and 5. We present the statistical properties of our sample in

[1] Although 2XMM J112328.0+052823 and 1E 1346+26.7 have Hβ FWHMs of 2000 and 2050 $\mathrm{km\,s^{-1}}$, respectively, they both have Hα FWHM of 1700 $\mathrm{km\,s^{-1}}$, and also share other NLS1s' spectral characteristics. Thus, they could both potentially be classified as NLS1s, making a total of 12.

Section 6. The summary and conclusions are given in Section 7. A flat universe model with a Hubble constant of $H_0 = 72\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, $\Omega_\mathrm{M} = 0.27$ and $\Omega_\Lambda = 0.73$ is adopted. In another paper, we will present our analysis of correlations between selected optical/UV emission features and the SED components, and discuss their physical implications (Jin et al., in preparation, hereinafter Paper II).

2 SAMPLE SELECTION AND DATA ASSEMBLY

To identify a sample of type 1 AGNs having both high-quality X-ray and optical spectra, we performed a cross-correlation between the 2XMMi catalogue and SDSS DR7 catalogue. We filtered the resulting large sample as described below. Our final sample consists of 51 type 1 AGNs including 12 NLS1s, all with high-quality optical and X-ray spectra and low reddening/absorption, and with Hβ linewidths ranging from 600 up to 13 000 $\mathrm{km\,s^{-1}}$. All the sources are listed in Table 1.

2.1 The cross-correlation between the 2XMMi and SDSS DR7 catalogues

The first step was to cross-correlate between the 2XMMi and SDSS DR7 catalogues. The 2XMMi catalogue contains 4117 *XMM–Newton* EPIC camera observations obtained between 03-02-2000 and 28-03-2008, and covering a sky area of $\sim$420 $\mathrm{deg^2}$. The SDSS DR7 is the seventh data release of the SDSS. The SDSS spectroscopic data have sky coverage of $\sim$8200 $\mathrm{deg^2}$, with spectra from 3800 to 9200 Å, and spectral resolution between 1800 and 2200 .

Our cross-correlation consisted of the following three steps:

(1) We first searched for all *XMM–Newton*/SDSS position pairs that lie within 20 arcsec of each other, resulting in 5341 such cases.

(2) For these 5341 unique X-ray sources, we imposed two further selection criteria: that source positions be separated by less than 3 arcsec or that sources be separated by no more than three times the *XMM–Newton* position uncertainty and no more than 7 arcsec. This filtering resulted in 3491 unique X-ray sources. The 3-arcsec separation is chosen because we want to include all possible *XMM–Newton*/SDSS pairs during these early filtering steps. From the 2XMMi and SDSS DR7 cross-correlation, there are 114 *XMM–Newton*/SDSS pairs whose separations are less than 3 arcsec, but are still nevertheless greater than three times the *XMM–Newton* position uncertainty. We included all of these pairs. The 7-arcsec separation upper limit mitigates spurious matches, especially for fainter objects and/or those located far off-axis.

(3) We selected only objects classified as extragalactic, giving a total of 3342 for further analysis.

2.2 Selection of type 1 Seyfert galaxies with high-quality spectra

Within these 3342 unique X-ray sources which satisfied all the above criteria, we applied further filtering to select only those type 1 AGNs having both high-quality optical and X-ray spectra. The five steps in the filtering were as follows:

(1) In order to obtain black hole mass estimates, we require Hβ and Hα emission lines to be measurable. Thus, we only selected sources with Hβ in emission (as indicated by the SDSS Hβ line models with at least 3σ significance and equivalent width EW >0) and redshift $z < 0.4$. This selection resulted in 802 unique X-ray sources, and 888 *XMM–Newton*/SDSS pairs (since some X-ray objects were matched with more than one SDSS spectrum).

(2) We then searched for the type 1 AGNs (including subtypes 1.0, 1.5, 1.8 and 1.9) which have a minimum of 2000 counts in at least one of the three EPIC cameras. Our search retrieved 96 such broad-line AGNs. We then inspected each of these *XMM–Newton*/SDSS pairs to confirm that all the matches were indeed genuine.

(3) From inspection of the SDSS spectra, we excluded 22 sources whose blueward part of the Hβ line showed strong reddening or low signal-to-noise ratio (S/N), which would distort the Hβ line profile. We also excluded one object, RBS 0992, because its SDSS spectrum did not show an Hβ line, due to a bad data gap. We ensured that the remaining 73 objects all had good Hβ line profiles.

(4) As a simple method to assess the spectral quality of the X-ray data, we used the *wabs*powerlaw* model in XSPEC v11.3.2 to fit the rest-frame 2–10 keV X-ray spectra of all 73 objects. The *error* command was used to estimate the 90 per cent confidence region for the *photon index* parameter. Based on the results, 16 objects with *photon index* uncertainties greater than 0.5 were thereby excluded, leaving 57 type 1 AGNs with relatively well constrained 2–10 keV spectra.

(5) By examining the 0.2–10 keV X-ray spectra, we excluded another six objects (i.e. IRAS F09159+2129, IRAS F12397+3333, PG 1114+445, PG 1307+085, PG 1309+355 and PG 1425+267) whose spectral shapes all showed clear evidence of an absorption edge at $\sim$0.7 keV (possibly originating from combined Fe I L-Shell and O VII K-Shell absorptions, Lee et al. 2001; Turner et al. 2004). This is a typical spectral signature of a warm absorber (e.g. Nandra & Pounds 1994; Crenshaw, Kraemer & George 2003). By removing such objects with complex X-ray spectra, our broad-band SED fitting is simplified. Our final sample contains 51 type 1 AGNs.

2.3 Characteristics of the sample

The sample-selection procedure described above ensures that every source in our AGN sample has both high-quality optical and X-ray spectra. In addition, a large fraction of the sample have simultaneous optical/UV photometric points from the OM monitor. Such high-quality data enable accurate spectral fitting. In the optical band, our sample is selected to have low reddening, since if present this would significantly modify the intrinsic continuum as well as the optical emission lines. This requirement reduces the complexity and uncertainty in our modelling of the intrinsic continuum, and also increases the overall quality of Hβ and Hα line profiles useful for estimating the black hole masses. Furthermore, low reddening is essential in the UV band. The inclusion of OM-UV photometric data observed simultaneously with the X-ray spectra provides a reliable link between these bands. This helps to reduce the fitting uncertainty of the SED resulting from optical and X-ray variability. Besides, all sources are well constrained in the 2–10 keV band, which is directly associated with the compact emitting region of the AGN. Our exclusion of objects with evidence of a warm absorber means that the 2–10 keV spectral index is likely to be intrinsic rather than hardened by absorption in the soft X-ray region.

In summary, compared with previous AGN samples used for broad-band SED modelling, the spectrally 'cleaner' nature of our sample should make the reconstructed broad-band SEDs more reliable. Consequently, the parameters derived from the broad-band spectral fitting should be more accurate. This may reveal new and potentially important broad-band correlations, which we will discuss in detail in Paper II.

Table 1. The type 1 Seyfert galaxy sample set.

ID	Common Name[a]	Redshift	2XMMi catalogue IAU name (2XMM[b])	*XMM–Newton* Observation date	SDSS DR7 MJD-plate-fibre	SDSS Observation date	EPIC Counts[c]
1	UM 269	0.308	J004319.7+005115	2002-01-04	51794-0393-407	2000-09-07	19 126
2	Mrk 1018	0.043	J020615.9−001730	2005-01-15	51812-0404-141	2000-09-25	2056
3	NVSS J030639	0.107	J030639.5+000343	2003-02-11	52205-0709-637	2001-10-23	35 651
4	2XMMi/DR7	0.145	J074601.2+280732	2001-04-26	52618-1059-399	2002-12-10	9679
5	2XMMi/DR7	0.358	J080608.0+244421	2001-10-26	52705-1265-410	2003-03-07	2912
6	HS 0810+5157	0.377	J081422.1+514839	2003-04-27	53297-1781-220	2004-10-19	4189
7	RBS 0769	0.160	J092246.9+512037	2005-10-08	52247-0766-614	2001-12-04	32 731
8	RBS 0770	0.033	J092342.9+225433*	2006-04-18	53727-2290-578	2005-12-23	104 028
9	Mrk 0110	0.035	J092512.8+521711	2004-11-15	52252-0767-418	2001-12-09	515 453
10	PG 0947+396	0.206	J095048.3+392650	2001-11-03	52765-1277-332	2003-05-06	58 555
11	2XMMi/DR7	0.373	J100025.2+015852	2003-12-10	52235-0501-277	2001-11-22	7187
12	2XMMi/DR7	0.206	J100523.9+410746	2004-04-20	52672-1217-010	2003-02-02	5437
13	PG 1004+130	0.241	J100726.0+124856	2003-05-04	53055-1744-630	2004-02-20	3781
14	RBS 0875	0.178	J103059.0+310255	2000-12-06	53440-1959-066	2005-03-11	69 434
15	KUG 1031+398	0.043	J103438.6+393828	2002-05-01	53002-1430-485	2003-12-29	63 891
16	PG 1048+342	0.160	J105143.8+335927	2002-05-13	53431-2025-637	2005-03-02	47 858
17	1RXS J111007	0.262	J111006.8+612522*	2006-11-25	52286-0774-600	2002-01-12	6147
18	PG 1115+407	0.155	J111830.2+402554	2002-05-17	53084-1440-204	2004-03-20	64 601
19	2XMMi/DR7	0.101	J112328.0+052823	2001-12-15	52376-0836-453	2002-04-12	10 098
20	RX J1140.1+0307	0.081	J114008.7+030710	2005-12-03	51994-0514-331	2001-03-26	35 616
21	PG 1202+281	0.165	J120442.1+275412	2002-05-30	53819-2226-585	2006-03-25	66 550
22	1AXG J121359+1404	0.154	J121356.1+140431	2001-06-15	53466-1765-058	2005-04-06	12 975
23	2E 1216+0700	0.080	J121930.9+064334	2002-12-18	53140-1625-134	2004-04-26	8028
24	1RXS J122019	0.286	J122018.4+064120	2002-07-05	53472-1626-292	2005-04-12	8338
25	LBQS 1228+1116	0.236	J123054.1+110011	2005-12-17	52731-1232-417	2003-04-02	165 823
26	2XMMi/DR7	0.304	J123126.4+105111	2005-12-17	52731-1232-452	2003-04-02	8816
27	Mrk 0771	0.064	J123203.6+200929	2005-07-09	54481-2613-342	2008-01-15	40 705
28	RX J1233.9+0747	0.371	J123356.1+074755	2004-06-05	53474-1628-394	2005-04-14	6041
29	RX J1236.0+2641	0.209	J123604.0+264135*	2006-06-24	53729-2236-255	2005-12-25	17 744
30	PG 1244+026	0.048	J124635.3+022209	2001-06-17	52024-0522-173	2001-04-25	8509
31	2XMMi/DR7	0.316	J125553.0+272405	2000-06-21	53823-2240-195	2006-03-26	7591
32	RBS 1201	0.091	J130022.1+282402	2004-06-06	53499-2011-114	2005-05-09	209 458
33	2XMMi/DR7	0.334	J132101.4+340658	2001-01-09	53851-2023-044	2006-04-26	4425
34	1RXS J132447	0.306	J132447.6+032431	2004-01-25	52342-0527-329	2002-03-09	6305
35	UM 602	0.237	J134113.9−005314	2005-06-28	51671-0299-133	2000-05-07	18 007
36	1E 1346+26.7	0.059	J134834.9+263109	2000-06-26	53848-2114-247	2006-04-23	71 985
37	PG 1352+183	0.151	J135435.6+180518	2002-07-20	54508-2756-228	2008-02-12	36 171
38	Mrk 0464	0.050	J135553.4+383428	2002-12-10	53460-2014-616	2005-03-31	13 974
39	1RXS J135724	0.106	J135724.5+652506	2005-04-04	51989-0497-014	2001-03-21	12 081
40	PG 1415+451	0.114	J141700.7+445606	2002-12-08	52728-1287-296	2003-03-30	55 786
41	PG 1427+480	0.221	J142943.0+474726	2002-05-31	53462-1673-108	2005-04-01	70 995
42	NGC 5683	0.037	J143452.4+483943	2002-12-09	52733-1047-300	2003-04-04	18 885
43	RBS 1423	0.208	J144414.6+063306	2005-02-11	53494-1829-464	2005-05-04	37 568
44	PG 1448+273	0.065	J145108.7+270926	2003-02-08	54208-2142-637	2007-04-18	134 532
45	PG 1512+370	0.371	J151443.0+365050	2002-08-25	53083-1353-580	2004-03-14	40 432
46	Q 1529+050	0.218	J153228.8+045358	2001-08-21	54563-1835-054	2008-04-07	10 952
47	1E 1556+27.4	0.090	J155829.4+271715	2002-09-10	52817-1391-093	2003-06-27	6995
48	Mrk 0493	0.031	J155909.6+350147	2003-01-16	53141-1417-078	2004-05-14	124 115
49	II Zw 177	0.081	J221918.5+120753	2001-06-07	52221-0736-049	2001-11-08	36 056
50	PG 2233+134	0.326	J223607.6+134355	2003-05-28	52520-0739-388	2002-09-03	7853
51	Mrk 0926	0.047	J230443.3−084111	2000-12-01	52258-0725-510	2001-12-15	59 513

[a]For some targets without well-known names, we simply use '2XMMi/DR7'.
[b]The full name should be '2XMM J...', but for those targets with an * symbol, the full names should be '2XMMi J...';
[c]The total counts in all three EPIC monitors, namely pn, MOS1 and MOS2, and there are at least 2000 counts in at least one of these three monitors.

2.4 Additional data

The 51 type 1 AGNs all have SDSS survey-quality spectra (flagged as 'sciencePrimary' in the SDSS catalogue), including three objects that have multiple SDSS spectra (i.e. NVSS J030639, 1RXS J111007 and Mrk 1018). In such cases, we adopt the SDSS spectrum which connects most smoothly with the OM data.

For each object, we used all available EPIC X-ray spectra (i.e. pn, MOS1 and MOS2) for the broad-band SED modelling, unless the spectrum had few counts and low S/N. We also searched through the *XMM–Newton*-OM SUSS catalogue for all data in the OM

bands (i.e. *V*, *B*, *U*, *UVW*2, *UVM*2 and *UVW*1), which are observed simultaneously with the corresponding EPIC spectrum. Of our 51 sources, we have 14 sources with SDSS optical spectra and *XMM–Newton* EPIC X-ray spectra, and 37 sources which in addition to these also have *XMM–Newton*-OM photometry.

2.5 OM data corrections and aperture effects

In the procedure of combining the SDSS spectra and OM data points, we identified that in some objects there is a clear discrepancy between these two data sets. The OM points often appear higher in the spectral plots (brighter) than is consistent from a smooth extrapolation of the SDSS spectral shape. In fewer cases, this discrepancy appears in the opposite sense, with the OM points apparently too low (fainter) (see Fig. 1 for some examples). This discrepancy may arise for several reasons, including a simple aperture effect. Compared to 3 arcsec diameter for the SDSS spectroscopy fibres, the OM monitor has a much larger aperture, that is, 12 and 35 arcsec diameters for the OM optical and OM UV filters, respectively (Talavera 2009). If the host galaxy is sufficiently extended, for example, in the case of RE J1034+396, the larger aperture of the OM would include more host galaxy emission than that in the SDSS spectrum (see also Section 5.3.1 for other possible reasons to account for this discrepancy). To investigate the aperture issue in more detail, we performed the following tests:

(1) We examined the combined SDSS and OM data plots, searching for those objects with excess OM flux compared with that expected from the extrapolated SDSS spectrum. We identified 27 such cases out of the 51 sources.

(2) Within this sample of 27 sources, we checked the catalogue flag for an extended source in each OM filter. We noted those flagged as an extended source in at least one OM filter. This yielded 13 sources out of the 27.

(3) We also extracted the SDSS CCD images for all 51 objects and visually checked whether they appeared extended. As a result, we included another four objects for which their SDSS CCD images show that their host galaxy is extended beyond the 3 arcsec diameter of the SDSS aperture. Either they were not flagged as extended sources in any OM filter or they did not have any OM optical data. For these 17 objects, an aperture effect could at least be partially responsible for an excess flux in the OM data.

(4) For these 17 objects, we downloaded all available OM image files. In each OM image, we applied a 6-arcsec-diameter aperture from which to extract the flux. We used the same-sized aperture placed on a blank region of the sky close to the object to estimate the background. The quoted point spread function (PSF) FWHM of the OM for the different filters are: *V*(1.35 arcsec), *B*(1.39 arcsec), *U*(1.55 arcsec), *UVW*1(2.0 arcsec), *UVM*2(1.8 arcsec) and *UVW*2(1.98 arcsec). Thus, in all cases, 6 arcsec is at least 3 × PSF FWHM. Thus, this aperture includes effectively all optical flux for a point source, and more than 90 per cent that from a UV point source detected by the OM.

Before subtracting the background flux from the source+background flux, we performed three count rate calibrations, according to the method described in the OM instrument document.[2] The first is the deadtime correction, required because for a small fraction of the exposure time, the CCD is in readout mode, and so cannot record events. The second calibration is for coincidence losses, which occur when more than 1 photon arrives on the CCD at the same location and within the same time-frame, so results in undercounting. The third calibration is for the OM time-sensitivity degradation correction. We performed these calibrations, according to the algorithms set out in the OM instrument document, separately for the background and source+background count rates. We then subtracted the background count rate from the source+background count rate to obtain the corrected source count rate.

Fig. 1 shows the OM data points before and after correction for aperture effects for the 17 objects. The reduced OM aperture does improve the alignment between the OM points and SDSS spectrum. This correction not only lowers the OM flux, but also changes the continuum shape defined by the OM points. Although the choice of an aperture smaller than 6 arcsec will lower the OM fluxes by a larger factor, it will also introduce uncertainties and systematics caused by the PSF. Therefore, we compromise by adopting a 6-arcsec-diameter aperture. In our subsequent SED modelling, we use the aperture-corrected OM data.

3 OPTICAL SPECTRAL MODELLING: THE EMISSION LINES

Our optical spectral modelling employs linked Hα and Hβ profile fitting and the complete optical spectral fitting. We wrote the code in IDL (Interactive Data Language) v6.2 to perform all the optical spectral fitting. The 'MPFITEXPR' program from the Markwardt IDL Library is incorporated within our code to perform the Levenberg–Marquardt least-squares algorithm used to obtain the best-fitting parameters. The SDSS spectra (stored in SDSS *spSpec* files) were extracted directly from the SDSS DR7 data archive and analysed in IDL using our code. A detailed description of our spectral modelling procedures is presented in the following subsections.

3.1 Profile fitting of the Hα, Hβ and [O III] λ5007 emission lines

Based on current AGN emission-line models, there are thought to be stratified regions emitting different lines. These regions are divided somewhat arbitrarily into a narrow-line region (NLR), a broad-line region (BLR) and possibly an intermediate-line region (ILR, e.g. Grupe et al. 1999; Hu et al. 2008; Mei, Yuan & Dong 2009; Zhu, Zhang & Tang 2009). Following previous studies, we use several separate Gaussian profiles representing each of these emitting regions to model the Balmer line profiles.

The Hα and Hβ line profiles each poses distinct difficulties for the spectral analysis. In the case of the Hβ line, the permitted Fe II emission features (which are often strong in NLS1s) and broad He II λ4686 line blend with the Hβ line, which can affect the determination of the underlying continuum and hence the Hβ line profile. For the Hα line, there is the problem of blending with the [N II] $\lambda\lambda$6584,6548 doublet, improper subtraction of which may distort Hα's intrinsic profile. Our approach therefore is to fit Hα and Hβ simultaneously using the same multi-Gaussian components. The assumed similarity between the intrinsic profiles of these two Balmer lines assists in deblending from other nearby emission lines, and should yield a more robust deconvolution for the separate components of their profile.

[2] URL: http://xmm2.esac.esa.int/docs/documents/CAL-TN-0019.ps.gz. Also see the XMM–Newton User Handbook: http://xmm.esac.esa.int/external/xmm_user_support/document-ation/uhb/index.html.

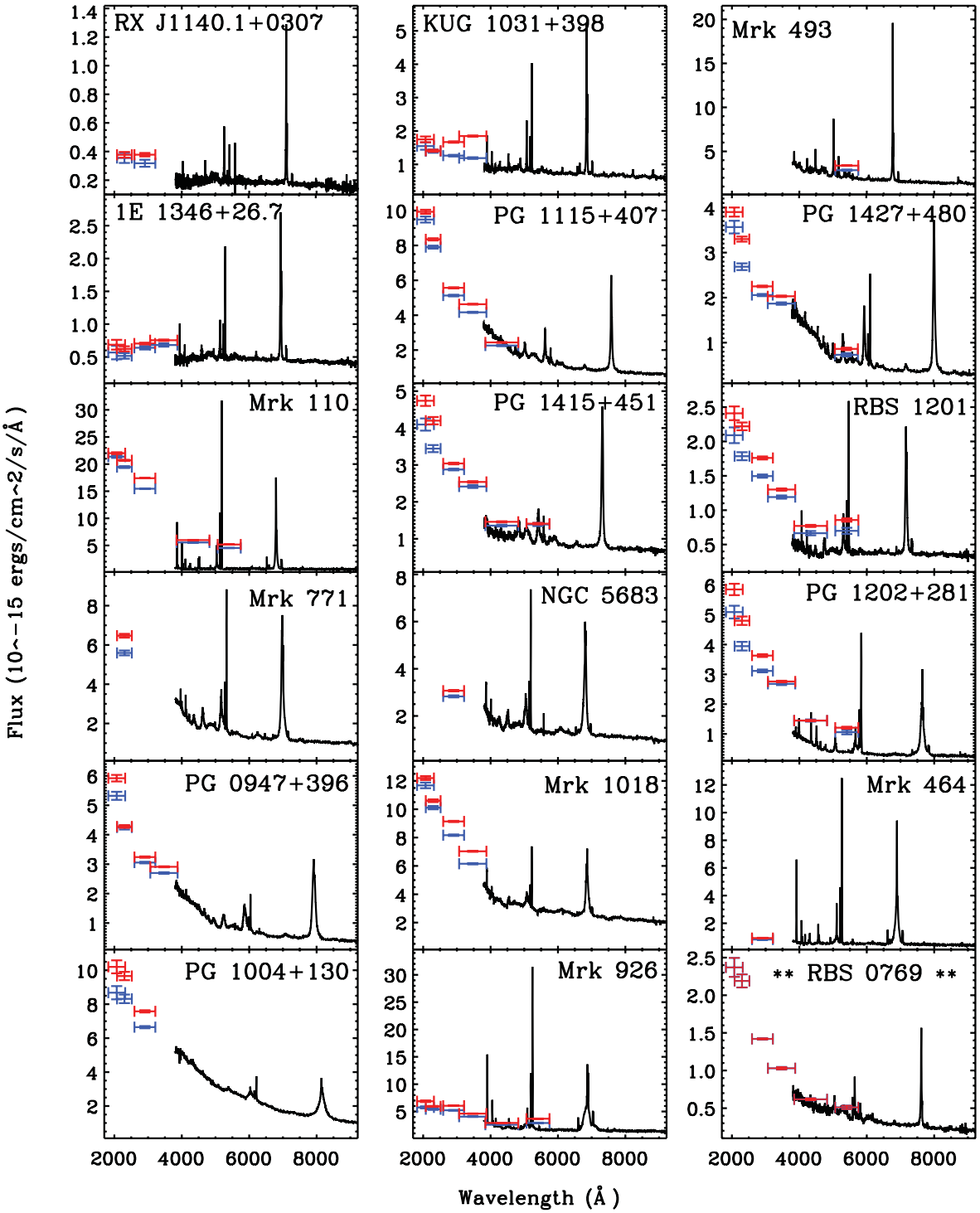

Figure 1. The aperture effect correction results for 17 extended sources in the sample. The point-like source RBS 0769 (last figure marked by **) is also shown for comparison. We overplot OM data points on to the SDSS spectrum. The red OM points are data obtained directly from the OM PPS files. The blue OM points are the corresponding data after applying a smaller 6-arcsec aperture to all OM filters, and applying appropriate OM corrections to the flux, for example, deadtime correction, coincidence loss correction and OM time-sensitivity degradation correction.

3.2 The Fe II problem

We use the theoretical Fe II model templates of Verner et al. (2009). These include 830 energy levels and 344 035 transitions between 2000 and 12 000 Å, totalling 1059 emission lines. The predicted Fe II emission depends on physical conditions such as microturbulence velocity and hardness of the radiation field, but we use the template which best matches the observed spectrum of I Zw 1 (Boroson & Green 1992; Véron-Cetty, Joly & Véron 2004), that is, the one with $n_{\rm H} = 10^{11}\,{\rm cm}^{-3}$, $v_{\rm turb} = 30\,{\rm km\,s}^{-1}$ and $F_{\rm ionizing} = 20.5\,{\rm cm}^{-2}\,{\rm s}^{-1}$. Detailed modelling of high-S/N spectra shows that the Fe II emission is often complex, with four major line systems in the case of I Zw

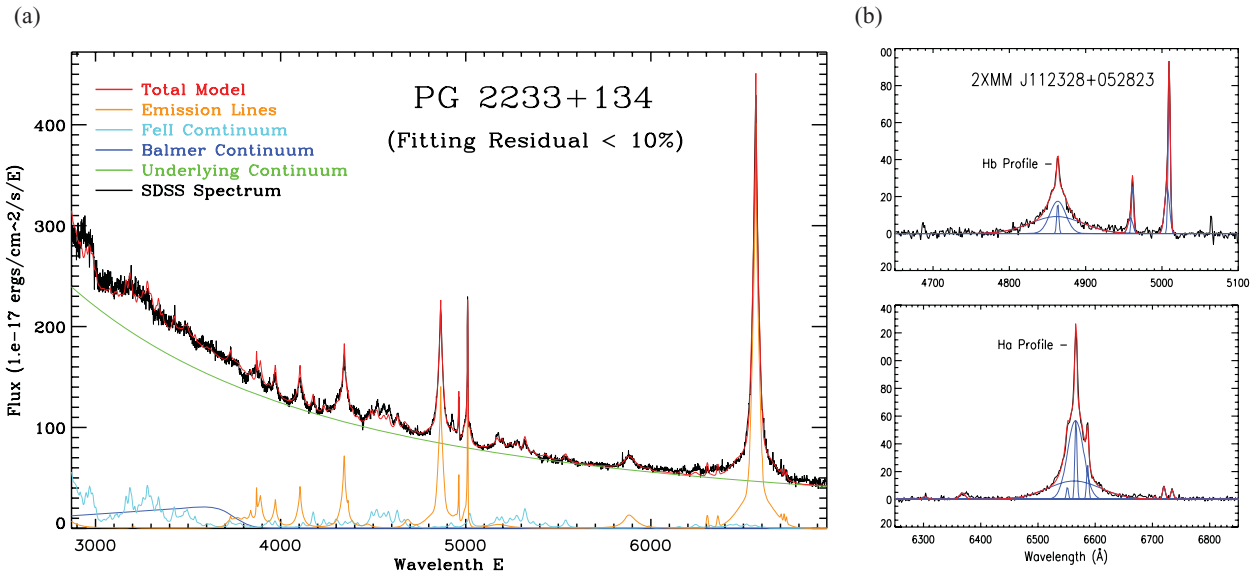

Figure 2. An example of results from SDSS spectrum fitting. Panel (a) shows a good fit for PG 2233+234. The black line is the observed spectrum and the red line is the total model spectrum. The green line represents the observed underlying continuum. The Balmer continuum (blue), Fe II emission (light blue) and other strong emission lines (orange) are shown underneath. Panel (b) shows an example of detailed line profile fitting to the Fe II-subtracted region around the Hβ (upper) and Hα lines (lower) including Hα, Hβ, [O III] $\lambda\lambda$5007, 4959 doublets, [N II] $\lambda\lambda$6585, 6548 doublets, Li λ6708, [S II] $\lambda\lambda$6717, 6733 doublets and [O I] $\lambda\lambda$6300, 6363 doublets. In our profile fitting, three Gaussian components are used for Hβ and Hα, two components for [O III] λ5007, and one Gaussian for all other lines. The various Gaussian profiles are shown in blue and the total model is shown in red.

1 (one broad-line system, two narrow high-excitation systems and one low-excitation system, Véron-Cetty et al. 2004; Zhou et al. 2006; Mei et al. 2009). However, for simplicity, we will assume only one velocity structure and convolve this template with a single Lorentzian profile.

We fit this to the actual Fe II emission-line features between 5100 and 5600 Å (no other strong emission lines lie in this wavelength range) of the de-redshifted SDSS spectra, leaving the FWHM of the Lorentzian and the normalization of Fe II as free parameters. The resulting best-fitting Fe II model to this restricted wavelength range was then extrapolated and subtracted from the entire SDSS spectrum. A major benefit from subtracting the Fe II features is that the profiles of the [O III] λ5007 lines no longer have apparent red wings. This is particularly important for the NLS1s, where the Fe II emission is often strong. After subtracting Fe II, we used either two or three Gaussian components (depending on the profile complexity) to fit the [O III] λ5007 line.

3.3 Deconvolution of the Balmer lines

After fitting the [O III] λ5007 line, we start to fit the Hα and Hβ line profiles simultaneously. Following previous studies, we consider a simplified picture in which the Balmer lines have three principal components, namely a narrow component (from the NLR), an intermediate component (from a transition region ILR between the NLR and BLR or from the inner edge of dusty torus, Zhu et al. 2009), and a broad component (from the BLR). The intermediate and broad components are both represented by a Gaussian profile, whereas the narrow component is assumed to be similar to that of [O III] λ5007. Since we do not know whether or not the Balmer decrements are the same in these different emitting zones, the relative strengths of different line components were not fixed, but their FWHM and relative velocity were both kept the same. The [O III] λ4959 line was set at one-third that of [O III] λ5007 from atomic physics. The [N II] $\lambda\lambda$6584,6548 line doublet was also fixed to the [O III] λ5007 line profile. For simplicity, the [S II] $\lambda\lambda$6733,6717 doublet, [O I] $\lambda\lambda$6300,6363 doublet and Li λ6708 were all fitted with a single Gaussian profile separately, because they are all relatively weak lines and do not severely blend with Balmer lines.

In order to separate the narrow component of the Balmer lines from the other components as accurately as possible, particularly for NLS1s and some broad-line objects which lack clear narrow-line profiles, we applied the following four different fitting methods:

(1) The profile of the narrow component is held the same as the entire [O III] λ5007 profile, and the normalization of each component in the Hα and Hβ lines is left as a free parameter.

(2) Only the central narrow component of the [O III] λ5007 profile is used to define the profile of the Balmer narrow component and of the [N II] $\lambda\lambda$6585,6550 doublet; the normalization of each component in the Hα and Hβ lines is a free parameter.

(3) The shape of the narrow component is held the same as the entire [O III] λ5007 profile, and also the normalization of the Hβ line narrow component is set to be 10 per cent of [O III] λ5007, this ratio being an average for the NLR in typical type 1 Seyfert galaxies (Osterbrock & Pogge 1985; Leighly 1999); all other components have their normalizations as free parameters.

(4) All conditions are the same as in method (3), except that the Balmer line narrow component and the [N II] $\lambda\lambda$6584,6548 doublet adopt the central narrow Gaussian component of the [O III] λ5007 line.

We applied each of the above fitting methods to every object in our sample, and then compared the results. For those objects with clear narrow components to their Balmer lines, we used the best-fitting result from methods (1) and (2). For the other objects whose narrow components were not clearly defined or even visible, we adopted methods (3) and (4), unless method (1) or (2) gave much better fitting results. Fig. 2(b) shows an example of our fitting. Results for

the whole sample are shown in Fig. A1 (Supporting Information with the online version of this paper).

After obtaining the best-fitting parameters, we used the intermediate and broad components to reconstruct the narrow-line subtracted Hβ line profile, and then measured the FWHM from this model. The rationale for using this method, instead of directly measuring the FWHM of the Hβ line from the data, is because for low-S/N line profiles, direct measurement of FWHM can lead to large uncertainties, whereas our profile models are not prone to localized noise in the data. The Hβ FWHM measurements for each of the 51 sources, after deconvolving using the instrumental resolution of 69 km s^{-1}, are listed in Table 3 (shown later).

4 OPTICAL SPECTRAL MODELLING

In order to obtain the underlying continuum, we must model the entire SDSS spectrum so that we can remove all the emission lines as well as the Balmer continuum and host galaxy contribution. As we are now concerned with the broad continuum shape, we choose to refit the Fe II spectrum across the entire SDSS range, rather than restricting the fit to the Hα and Hβ line regions as discussed in the previous section.

Fig. 2 shows an illustrative example of our optical spectral fitting, and the results for each of the 51 sources are presented in Fig. A1. In the following subsections, we give further details of the components that make up these modelled spectra.

4.1 Emission lines including Fe II

We use the models for [O III], Hα and Hβ as derived above. We add to this a series of higher order Balmer lines: from 5→2 (Hγ) to 15→2. We fix the line profile of these to that of Hβ up to 9→2, then simply use a single Lorentzian profile for the rest weak higher order Balmer lines. We fix the line ratios for each Balmer line using the values in Osterbrock (1989), table 4.2, with T_e between 10 000 and 20 000 K. We similarly use a single Lorentzian to model the series of helium lines (He I λ3187, He I λ3889, He I λ4471, He I λ5876, He II λ3204, He II λ4686) and some other emission lines (Mg II λ2798, [Ne III] λλ3346,4326, [O II] λλ3727,3730, [O I] λλ6302,6366, [N II] λλ6548,6584, Li λ6708, [S II] λλ6717,6733).

We use the same model for the Fe II emission as described in Section 3.1. However, we now fit this to the entire SDSS wavelength range, rather than restricting the fit to 5100–5600 Å.

4.2 The Balmer continuum

Another potentially significant contribution at shorter wavelengths is from the Balmer continuum. Canfield & Puetter (1981) and Kwan & Krolik (1981) predicted the optical depth at the Balmer continuum edge to be less than 1. We use equation (1) to model the Balmer continuum under the assumptions of the optically-thin case and a single-temperature electron population (also see Grandi 1982; Wills et al. 1985):

$$F_v^{\rm BC} = F_v^{\rm BE} {\rm e}^{-h(v-v_{\rm BE})/(kT_{\rm e})} \quad (v \geq v_{\rm BE}), \tag{1}$$

where $F_v^{\rm BE}$ is the flux at the Balmer edge, $v_{\rm BE}$ corresponds to the Balmer-edge frequency at 3646 Å, $T_{\rm e}$ is the electron temperature, h is the Planck constant and k is Boltzmann's constant. This Balmer continuum equation is then convolved with a Gaussian profile to represent the real Balmer bump in SDSS spectra.

There are several parameters that may slightly modify or significantly change the shape of the Balmer continuum. It is already seen that the electron temperature $T_{\rm e}$ appearing in equation (1) and the optical depth can both change the Balmer continuum shape, but there are additional important factors. Any intrinsic velocity dispersion will Doppler-broaden all the hydrogen emission features. Therefore, a better description of the Balmer continuum can be obtained by convolving equation (1) with a Gaussian profile, whose FWHM is determined by the linewidth of Hβ (or other broad lines), as shown by equation (2), where $G(x)$ represents a Gaussian profile with a specific FWHM:

$$F_\lambda^{\rm BC} = F_\lambda^{\rm BE} {\rm e}^{hc/(\lambda_{\rm BE}kT_{\rm e})} \int_0^{+\infty} {\rm e}^{-hc/(\lambda kT_{\rm e})} G(\lambda_1 - \lambda) {\rm d}\lambda_1. \tag{2}$$

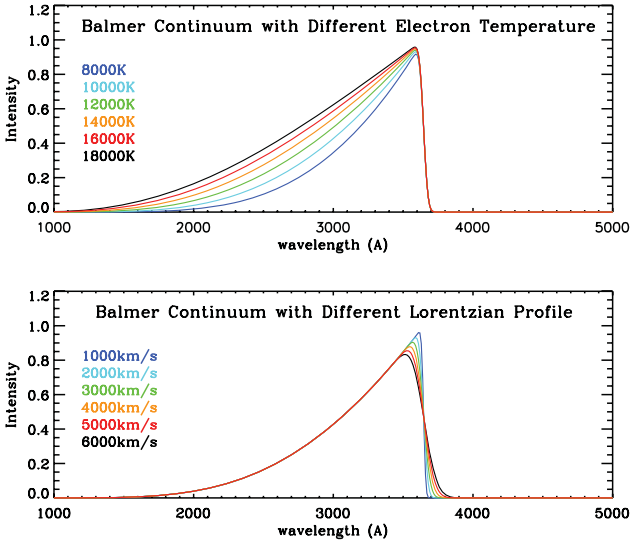

Figure 3. The Balmer continuum models of Grandi (1982). The upper panel shows the dependence of the model on electron temperature. The lower panel shows the dependence of the model on the FWHM of the convolved Lorentzian profile.

Fig. 3 shows how the Balmer continuum's shape depends on the electron temperature and velocity broadening in equation (2). The electron temperature modifies the decrease in the Balmer continuum towards shorter wavelengths, but has little effect on the broadening of the Balmer photorecombination (BPR) edge. On the contrary, velocity broadening mainly affects the shape of the BPR edge, but the emission longwards of 3646 Å is still very weak compared to the emission bluewards of the BPR edge, that is, the BPR edge is still sharp.

We initially applied equation (2) to fit the Balmer continuum bump below 4000 Å in the SDSS spectra. We assumed the velocity profile for the convolution was a Gaussian with its FWHM determined from the Hβ line profile, and the wavelength of the position of the BPR edge was taken as the laboratory wavelength of 3646 Å. However, this model did not provide an acceptable fit (see e.g. the model shown by the blue line in Fig. 4). It appears that the observed spectrum requires a model with either a more extended wing redwards of the BPR edge or a BPR edge that shifts to longer wavelength than 3646 Å. However, additional velocity broadening should affect both the Balmer continuum and Balmer emission lines equally, as they are produced from the same material, although the multiple components present in the line make this difficult to constrain.

One way the wavelength of the edge may be shifted without affecting the lines is via density (collisional, or Stark) broadening (e.g. Pigarov, Terry & Lipschultz 1998). Multiple collisions disturb the outer energy levels, leading to an effective $n_{\rm max}$ for the highest

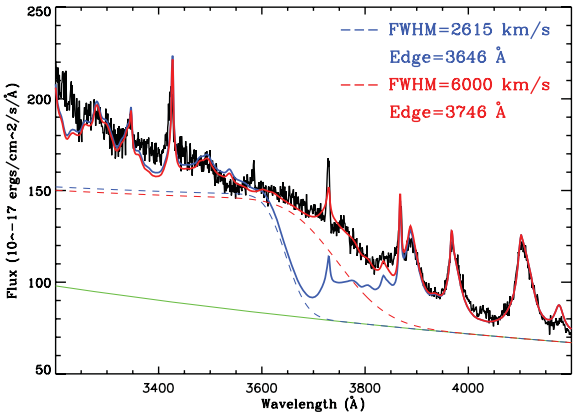

Figure 4. An expanded view of the region around the BPR edge in PG 1427+480. The blue and dashed lines represent the Balmer continuum model superposed on the underlying disc continuum (green solid line) using standard parameters (blue dashed line), and also a set of best-fitting parameters (red dashed line). The red and blue solid lines are models of the total optical spectrum, including the corresponding Balmer continuum components plus other components described in the text. The observed spectrum is shown in black.

bound level $\ll\infty$, that is, lowering the effective ionization potential. We set the edge position and the FWHM as free parameters, and let the observed spectral shape determine their best-fitting values. The red line shown in Fig. 4 represents a good fit, obtained with FWHM of 6000 km s^{-1} and the BPR edge wavelength of 3746 Å, which implies $n_{\rm max} \sim 12$. The theoretical $n_{\rm max}$ can be determined by the plasma density $N_{\rm e}$ and temperature $T_{\rm e}$ as $n_{\rm max} = 2 \times 10^4 (T_{\rm e}/N_{\rm e})^{1/4}$ (Mihalas 1978), so for a typical temperature of 10^4–10^5 K, the required density is 7×10^{16}–7×10^{17} cm^{-3}. Such high density is not generally associated with the BLR clouds, and may give support to models where the low-ionization BLR is from the illuminated accretion disc (e.g. Collin-Souffrin & Dumont 1990). However, any reliable estimation of the density would require more accurate subtraction of other optical components such as the Fe II line blends and many other non-hydrogen emission lines, which is not the focus of this paper. None the less, this remains an interesting problem which is worthy of further study.

Yet another issue in modelling the Balmer continuum is how to quantify the total intensity of this continuum component, especially when there is limited spectral coverage below 4000 Å, which makes it difficult to define the overall shape. The theoretical flux ratio between the Balmer continuum and the Hβ line under case B conditions can be expressed by equation (3) (Wills et al. 1985):

$$I(\mathrm{Bac})/I(\mathrm{H}\beta) = 3.95 T_4^{0.4}, \quad (3)$$

but other theoretical calculations of photoionization models show that by varying the Balmer optical depth, electron temperature and electron number density, this can result in very different values of I(Bac)/I(Hβ). For example, Canfield & Puetter's (1981) calculation resulted in a I(Bac)/Hα range of 0.05–10, Kwan & Krolik (1981) suggested I(Bac)/I(Hβ) = 1.6–15, and other theoretical work also confirmed a large range in flux ratios (Puetter & Levan 1982; Kwan 1984; Hubbard & Puetter 1985). The observed ranges in I(Bac)/I(Hβ) are also large. Canfield & Puetter (1981) showed an observed range of 0.5–3 for I(Bac)/I(Hα). Wills et al. (1985) observed nine intermediate-redshift quasi-stellar objects whose I(Bac)/I(Hβ) ranges from 4.65 to 9.5. Thus, we were unable to constrain the intensity of the whole Balmer continuum by using a standard flux ratio fixed to the other Balmer emission lines. As a result, we must rely on the shape of the observed Balmer bump, and then adopt the model's best-fitting parameters.

However, this limitation in defining the Balmer bump introduces uncertainties in modelling the underlying continuum, because over-subtraction of the Balmer bump will depress the slope of the remaining underlying continuum, and vice versa. In the course of the broad-band SED fitting described in Section 5, we found that the temperature of accretion disc (determined by black hole mass) is sensitive to the slope of optical continuum, unless the continuum slope is in the opposite sense to that of the accretion disc model and thus cannot be fitted, or there are OM points providing stronger constraints. We also found that a flatter optical continuum may lead to a lower best-fitting black hole mass, although this also depends on other factors. Therefore, the subtraction of the Balmer continuum can have an impact on the modelling of broad-band SED and the best-fitting black hole mass. The influence of this depends on the relative importance of other SED restrictions. This is the reason why the Balmer continuum must be carefully modelled and subtracted.

4.3 The intrinsic underlying continuum

Our basic assumption is that the residual optical spectrum, after subtraction of the Balmer continuum, Fe II emission and other emission lines mentioned previously, arises mainly from the accretion disc emission. As a reasonable approximation over a limited wavelength range, we use a power law of the following form to fit the underlying continuum:

$$F(\lambda) = C_1 (\lambda/5100\,\text{Å})^{-C_2}. \quad (4)$$

The power-law approximation for the underlying optical disc continuum is also widely adopted in previous and recent AGN optical spectral studies (e.g. Grandi 1982; Tsuzuki et al. 2006; Zhou et al. 2006; Landt et al. 2011).

We model the dust reddening using Seaton's (1979) 1100–10 000 Å reddening curve, and we apply this to the overall model, that is, emission lines, the Balmer continuum and disc continuum. There are also other reddening curves available such as Fitzpatrick (1986) for the Large Magellanic Cloud, Prévot et al. (1984) and Bouchet et al. (1985) for the Small Magellanic Cloud and Calzetti et al. (2000) for starburst galaxies, but over the wavelength range of 2500–10 000 Å, the difference between these reddening curves is small, except for Calzetti et al.'s (2000) curve which is appropriate for starburst galaxies, and is thus not applicable for our AGN sample.

4.4 The host galaxy contribution

Many previous studies on AGN's optical/infrared spectra have adopted a power law as a reasonable approximation for the accretion disc continuum bluewards of 1μm (e.g. Mei et al. 2009; Bian & Huang 2010), but these studies also needed to include additional contributions from the host galaxy and emission from the dusty torus to account for the extra continuum emission at long wavelengths of the optical spectrum (e.g. Kinney et al. 1996; Mannucci et al. 2001; Landt et al. 2011). In our work, we have also identified an inconsistency between the 3000–8000 Å spectral shape and a single power-law shape (i.e. the flat optical spectrum problem discussed in Section 5.3.2). The blue end of the optical spectrum, presumed to arise from a standard accretion disc, often shows a steeper spectral slope than the red end.

However, in our sample, we found evidence suggesting only a weak, if any, contribution from the host galaxy. For example, the

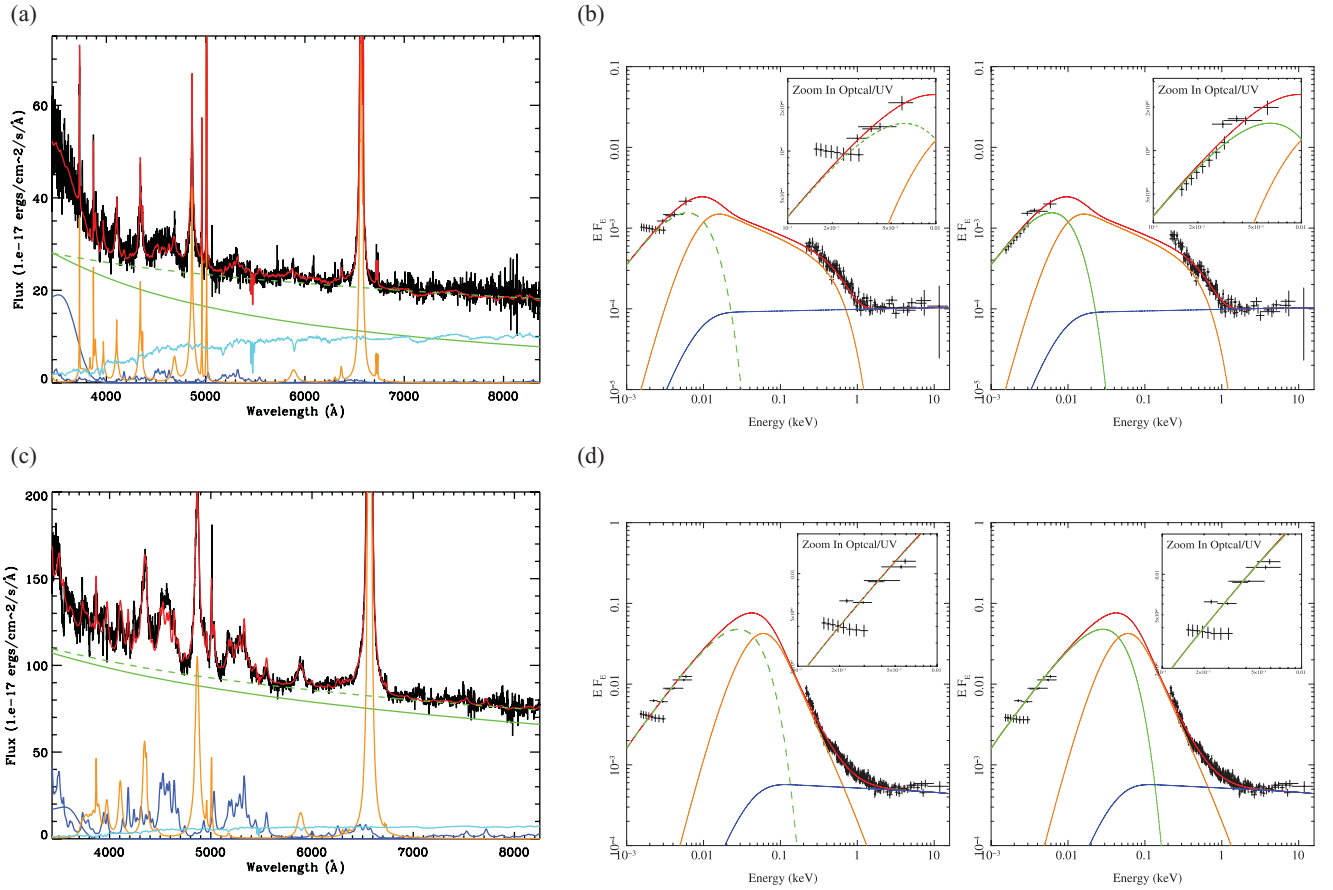

Figure 5. A comparison between the results of two subtractions of host galaxy contribution. 2XMM J112328+052823 (panels a and b) shows an underlying continuum that more closely resembles a disc continuum (solid green line in panel a) after modelling and subtracting the host galaxy contribution (light blue spectrum in panel a). The left-hand plot of panel (b) shows the original broad-band SED fitting without subtracting the host galaxy contribution. The dashed green line shows the modelled accretion disc emission in the best-fitting SED. The insets show a magnification of the fit in the optical/UV region, where a big discrepancy exists between the SDSS data and best-fitting SED model. The right-hand plot of panel (b) is the new SED fit using the new underlying disc continuum (shown as the solid green line in panel a) after subtracting the host galaxy contribution. The new fit is improved in the optical region compared with the previous results in the left-hand plot of panel (b). In contrast to the above example, PG 1415+451 (panels c and d) has little host galaxy contribution in the SDSS optical spectrum (see the light blue component in panel c), and its broad-band SED fitting in the optical region remains poor regardless of the amount of host galaxy subtraction applied (see the two plots in panel d). The spectral template for elliptical galaxies in Kinney et al. (1996) was used in both cases since their host galaxies both have elliptical morphologies in the SDSS image.

optical spectra of our sample do not show the strong curvature characteristic of the presence of a stellar component in a host galaxy. Furthermore, the good-quality optical spectra do not exhibit stellar absorption features (see Section 5.3.2 and Fig. 5). In fact, the 3-arcsec-diameter fibre used to obtain the SDSS spectra also helps to reduce the contribution of stellar emission from a host galaxy, particularly for nearby sources in our sample such as KUG 1031+398. This evidence argues against the possibility that the red optical continuum is primarily dominated by host galaxy emission. In fact, it is possible that the observed additional component arises due to emission from the outer regions of a standard accretion disc (e.g. Soria & Puchnarewicz 2002; Collin & Kawaguchi 2004; Hao et al. 2010). The existence of such an additional red optical continuum component reduces the consistency of a power-law fit to the optical spectra.

4.5 The optical spectrum fitting

Our optical spectral fitting is performed only for data bluewards of 7000 Å. The choice to truncate the model at 7000 Å is made for several reasons. We wish to include an Hα line in the spectral-fitting range, and the broad wing of the Hα profile sometimes extends to ∼7000 Å (e.g. PG 1352+183, RBS 1423, Mrk 926). There are some objects whose SDSS spectra extend only to ∼6700 Å (e.g. 2XMM J080608.0+244421, HS 0810+5157, 2XMM J100025.2+015852). The choice of 7000Å, rather than a longer wavelength, is to maintain consistency of optical spectral fitting for the whole sample. The final reason concerns an aspect of the power-law fitting. We found that in some objects (e.g. PG 1115+407, LBQS 1228+1116, PG 1352+183), a flat slope power law underpredicts the observed emission at ∼7000 Å. Therefore, if we include longer wavelengths than 7000 Å, our power-law fitting for the standard accretion disc continuum towards the blue optical spectra would be biased by other continuum emission at these longer wavelengths, and so affects the broad-band SED fitting. Consequently, we chose to truncate our optical spectral fitting at 7000 Å.

However, we still cannot be sure that the underlying continuum is totally free from other non-disc continuum components. Thus, after completing the fitting procedure, we then checked the spectral fitting status within two narrow wavebands, that is, 4400–4800 and

5100–5600 Å. Emission features if present in these two wavebands are mainly from Fe II emission, and the underlying continua of these two wavebands should be totally dominated by the accretion disc emission. Assuming that the Fe II emission lines within these two wavebands have similar relative intensity ratios to those in the Fe II template described in Section 3.2, the best-fitting underlying power-law plus Fe II emission model should have good fitting status in both these two wavebands. In general, the best-fitting model derived from the full optical spectrum fit also gives reasonably good fitting status in both these two narrow wavebands. However, in some cases, the model overpredicted the flux in 5100–5600 Å but underpredicted the flux in 4400–4800 Å, so that we should slightly increase the slope of power law to produce better spectral fitting in these two wavebands. We adopted these parameter values in preference to those directly from the full spectrum fit, as they should be more immune to problems such as host galaxy or hot dust contamination.

5 THE BROAD-BAND SED MODELLING

5.1 Data preparation

For each object, we extracted the original data files and the pipeline products (PPS) from the XMM–Newton Science Archive.[3] In the following data-reduction process, tasks from the XMM–Newton Science Operations Centre SAS v7.1.0 (Science Analysis System) software were used. First, the `EPCHAIN/EMCHAIN` task was used to extract events unless the events files had already been extracted for each exposure by PPS. Then, the `ESPFILT` task was used to define background good time intervals that are free from flares. In each available EPIC image, a 45-arcsec-radius circle was used to extract the source region, and an annulus centred on the source with an inner and outer radius of 60 and 120 arcsec, respectively, was used to define the background region. For other sources listed in the region files of PPS that are included in these regions, these were subtracted using the default radii generated by PPS, which scaled with the source brightness. Then, the GIT filter, source and background region filters were applied to the corresponding events files to produce a set of source and background events files. We only accepted photons with quality flag = 0 and pattern 0–4. The `EPATPLOT` task was then used to check for pile-up effects. When pile-up was detected, an annulus with an inner and outer radius of 12 and 45 arcsec, respectively, was used instead of the previous 45-arcsec-radius circle to define the source region. Then, source events files were reproduced using the new source region filter. Source and background spectra were extracted from these events files for each available EPIC exposure. The task `RMFGEN/ARFGEN` was used to produce response matrices and auxiliary files for the source spectra. These final spectra were grouped with a minimum of 25 counts per bin using the GRPPHA v3.0.1 tool for spectral fitting in XSPEC v11.3.2. To prepare the OM data, the *om_filter_default.pi* file and all response files for the *V*, *B*, *U*, *UVW*1, *UVM*2 and *UVW*2 filters were downloaded from the OM response file directory in the HEASARC Archive.[4] We then checked the OM source list file for each object to see if there were any available OM count rates. Each count rate and its associated error were entered into the *om_filter_default.pi* file and then combined with the response file of the corresponding OM filter, again by using the GRPPHA tool to produce OM data that could be used in XSPEC.

Finally, the *XMM–Newton* EPIC spectra were combined with the aperture-corrected OM photometric points, and the optical continuum points produced from the optical underlying continuum (obtained from the full optical spectrum fitting) using the FLX2XSP tool. From these data, we constructed a broad-band nuclear SED of each AGN. There is a ubiquitous data gap in the far-UV region which is due to photoelectric absorption by Galactic gas. Unfortunately, in most cases of low-redshift AGNs, their intrinsic SED also peaks in this very UV region, and so this unobservable energy band often conceals a large portion of the bolometric luminosity. In order to account for this, and to estimate the bolometric luminosity, we fit the X-ray and UV/optical continua all together using a new broad-band SED model (Done et al. 2011, XSPEC model: *optxagn*). We then calculate the bolometric luminosity by summing up the integrated emission using the best-fitting parameters obtained for each continuum component.

5.2 The broad-band SED model

A standard interpretation of the broad-band SED is that the emission is dominated by a multi-temperature accretion disc component which peaks in the UV (e.g. Gierliński et al. 1999, XSPEC model: *diskpn*). This produces the seed photons for Compton upscattering by a hot, optically-thin electron population within a corona situated above the disc, resulting in a power-law component above 2 keV (e.g. Haardt & Maraschi 1991; Zdziarski, Poutanen & Johnson 2000, XSPEC model: *bknpl*). However, the X-ray data clearly show that there is yet another component which rises below 1 keV in almost all high-mass-accretion-rate AGNs. The ubiquity of this component can be seen, for example, in the compilation of AGN SEDs presented in Middleton et al. (2007), and two of the strongest cases are the NLS1s RE J1034+396 (Casebeer, Leighly & Baron 2006; Middleton et al. 2009) and RX J0136.9−3510 (Jin et al. 2009). The origin of this so-called soft X-ray excess is still unclear (e.g. Gierliński & Done 2004; Crummy et al. 2006; Turner et al. 2007; Miller, Turner & Reeves 2008), and so some previous broad-band SED modelling studies have explicitly excluded data below 1 keV. An obvious consequence is that in such studies a soft-excess component cannot influence the models, so making it possible to fit the data using just a disc and (broken) power-law continuum (VF07; Vasudevan & Fabian 2009). However, in our present study, we include all of the data, and so we require a self-consistent model which incorporates this soft component.

Whatever the true origin of the soft X-ray excess, the simplest model which can phenomenologically fit its shape is the optically-thick, low-temperature thermal Comptonization model (`compTT`). However, the observed data are used to constrain the three separate components `diskpn + compTT + bknpl`, which is generally problematic, given the gap in spectral coverage between the UV and soft X-ray regions caused by interstellar absorption. Thus, instead, we combine these three components together using a local model in XSPEC, assuming that they are all ultimately powered by gravitational energy released in accretion. A complete description of this model can be found on the XSPEC website[5] and is also given in Done et al. (2011). It is in essence a faster version of the models recently applied to black hole binary spectra observed close to their Eddington limit (Done & Kubota 2006) and to the (possibly super-Eddington) ultraluminous X-ray sources (Gladstone, Roberts & Done 2009; Middleton & Done 2010); thus, this model is more appropriate for fitting a medium-sized sample of objects.

[3] http://xmm.esac.esa.int/external/xmm_data_acc/xsa/index.shtml.
[4] http://heasarc.gsfc.nasa.gov/FTP/xmm/data/responses/om/

A comprehensive comparison with the model of Done & Kubota (2006) is given in Done et al. (2011). To make this paper self-contained, we give a brief synopsis of the model. We assume that the gravitational energy released in the disc at each radius is emitted as a blackbody only down to a given radius, $R_{\rm corona}$. Below this radius, we further assume that the energy can no longer be completely thermalized, and is distributed between the soft-excess component and the high-energy tail. Thus, the model includes all three components which are known to contribute to AGN SEDs in a self-consistent way. As such it represents an improvement on the fits in VF07 in several respects, by including the soft excess and by requiring energy conservation, and it improves on Done & Kubota (2006) by including the power-law tail.

In our SED fitting, the optical/UV data constrain the mass-accretion rate through the outer disc, provided we have an estimate of the black hole mass. We constrain this by our analysis of the Hβ emission-line profile. The main difference from previous studies based on non-reverberation samples is that we do not directly use the FWHM of the Hβ profile to derive the black hole mass. Rather, we use the FWHM of the intermediate- and broad-line components determined from the emission-line fitting results presented in Section 3.1. These are then used in equation (5) (Woo & Urry 2002, and references therein) to derive the black hole mass limits required for the SED fitting:

$$M_{\rm BH} = 4.817\times \left[\frac{\lambda L_{\lambda}(5100\ \text{Å})}{10^{44}\ \text{erg s}^{-1}}\right]^{0.7} \text{FWHM}^2, \quad (5)$$

where $L_{\lambda}(5100\ \text{Å})$ is measured directly from the SDSS spectra. The rms difference between the black hole masses from this equation and from the reverberation mapping study is $\sim$0.5 dex. Thus, we also adopted any best-fitting values that fell below the original lower limit (which was set by the FWHM of the intermediate component) by less than 0.5 dex. With this method, the best-fitting black hole mass found by SED fitting is always consistent with the prediction from the Hβ profile. Section 6.5 discusses the differences between the best-fitting black hole masses and those estimated using other methods.

Once the black hole mass is constrained, the optical data then set the mass accretion rate $\dot{M}$, and hence the total energy available is determined by the accretion efficiency. We assume a stress-free (Novikov–Thorne) emissivity for a Schwarzschild black hole, that is, an overall efficiency of 0.057 for $R_{\rm in} = 6R_{\rm g}$. Thus, the total luminosity of the soft excess and power law is $0.057\dot{M}c^2(1-R_{\rm in}/R_{\rm corona})$. This constrains the model in the unobservable extreme-UV region, with the input free parameter $R_{\rm corona}$ setting the model output of the luminosity ratio between the standard disc emission and Comptonization components. The upper limit of $R_{\rm corona}$ is set to be $100R_{\rm g}$, which corresponds to 81 per cent of the released accretion disc energy. This upper limit is based on the requirement that the seed photons should be upscattered (Done et al. 2011). We assume that both the Comptonization components scatter seed photons from the accretion disc with a temperature corresponding to $R_{\rm corona}$. The other model input parameters are the temperature ($kT_{\rm e}$) and optical depth (τ) of the soft-Comptonization component which are determined by the shape of the soft X-ray excess and the spectral index (Γ) of the hard X-ray Comptonization that produces the 2–10 keV power law, with electron temperature fixed at 100 keV . The model output $f_{\rm pl}$ represents the fraction of the non-thermalized accretion energy (i.e. given by the luminosity originating from the region of $R_{\rm corona}$ to $R_{\rm in}$), which is emitted in the hard X-ray Comptonization.

We also included two sets of corrections for attenuation (*reddening-wabs*) to account for the line-of-sight Galactic absorption and for the absorption intrinsic to each source, the latter being redshifted (*zred* and *zwabs* in XSPEC). The Galactic H I column density is fixed at the value taken from Kalberla et al. (2005), but the intrinsic H I column density is left as a free parameter. The standard dust-to-gas conversion formula of $E(B-V) = 1.7\times10^{-22}N_{\rm H}$ (Bessell 1991) is used for both Galactic and intrinsic reddening. We set the initial value of the power-law photon index to be that of the photon index in the 2–10 keV energy band, but it can vary during the fitting process. However, we set an upper limit of 2.2 for the power-law photon index, not only because the photon index is $<$2.2 for the majority of type 1 AGNs (Middleton et al. 2007), but also because otherwise the much higher S/N in the soft excess in some observed spectra can artificially steepen the hard X-ray power law and result in unphysical best-fitting models.

All free parameters used in the broad-band SED fitting are listed in Table 2. For completeness, we also explicitly calculate the fraction of the total luminosity carried by each component of the model (i.e. disc: $f_{\rm d}$; soft Comptonization: $f_{\rm c}$; hard X-ray Comptonization: $f_{\rm p}$) from the model-fitting parameters $R_{\rm cor}$ and $f_{\rm pl}$ (see Table 2).[5] Table 3 lists the important characteristic parameters. The main uncertainty in these parameters, especially the black hole mass, is dominated by other systematic uncertainties introduced by the observational data, model assumptions (e.g. the assumption of a non-spinning black hole and the inclination dependence of the disc emission) and the analysis methods involved. Therefore, the parameter-fitting uncertainties which are often less than 10 per cent are not significant in comparison, and thus are not listed. The statistical properties of these parameters are discussed in Section 6.

5.3 Problems in SED fitting

We further discuss two problems we encountered during the fitting procedure in the following subsections. The first problem is the discrepancy between the OM and SDSS continuum points (mentioned in Section 2.5). The second problem is that of the observed flat optical continuum, whose shape cannot be accounted for in our SED model (mentioned in Section 4.4).

5.3.1 The discrepancy between the OM photometry and the SDSS continuum

There remains a significant discrepancy between many of the OM and SDSS continuum points, even after applying the aperture correction discussed in Section 2.2 (see Fig. A1). The OM points often appear above (brighter) the extrapolation of the SDSS continuum to the OM wavelengths. We identify the following three possible reasons for this discrepancy:

(1) *Remaining aperture effects*. There is an aperture difference between the SDSS fibres (3 arcsec diameter) and the OM apertures we used (6 arcsec diameter). Clearly, the OM points will still include more host galaxy starlight than the SDSS points, and so will appear above the SDSS spectrum.

(2) *Contamination from emission lines*. The wavelength ranges for each OM filter (over which the effective transmission is greater

[5] A full description of the model parameters can be found on the XSPEC web page: http://heasarc.nasa.gov/xanadu/xspec/models/optxagn.html

Table 2. Broad-band SED-fitting parameters and model outputs ($L_{\rm bol}, f_{\rm d}, f_{\rm c}, f_{\rm p}$). ID: object number, the same as Table 1; $N_{\rm H,gal}$ and $N_{\rm H,int}$: the fixed Galactic and free intrinsic neutral hydrogen column densities in 10^{20} cm^{-2}; $\Gamma_{\rm pow}$: the power-law component's slope in the SED fitting, (*) denotes the objects whose power-law slopes hit the uplimit of 2.2 and were fixed there; $f_{\rm pl}$: the fraction of power-law component in the total reprocessed disc emission; $R_{\rm cor}$: corona (truncation) radius in units of gravitational radii ($r_{\rm g}$) within which all disc emission is reprocessed into the Comptonization and power-law components; $T_{\rm e}$: the temperature of the Compton upscattering electron population; τ: the optical depth of the Comptonization component; $\log(M_{\rm BH})$: the best-fitting black hole mass; $\log(\dot{M})$: the total mass accretion rate; $L_{\rm bol}$: the bolometric luminosity integrated from 0.001 to 100 keV; $f_{\rm d}, f_{\rm c}, f_{\rm p}$: the luminosity fractions of disc emission, soft-Comptonization and hard-X-ray-Compotonization components in the bolometric luminosity; and χ^2: the reduced χ^2 of the broad-band SED fitting.

ID	$N_{\rm H,gal}$ ($\times10^{20}$)	$N_{\rm H,int}$ ($\times10^{20}$)	$\Gamma_{\rm pow}$	$f_{\rm pl}$	$R_{\rm cor}$ ($r_{\rm g}$)	$T_{\rm e}$ (keV)	τ	$\log(M_{\rm BH})$ ($M_\odot$)	$\log(\dot{M})$ (g s^{-1})	$L_{\rm bol}$ ($\times10^{44}$)	$f_{\rm d}$	$f_{\rm c}$	$f_{\rm p}$	χ^2 reduced
1	1.79	0.00	1.71	0.69	100.	0.262	17.2	8.61	26.06	58.9	0.19	0.25	0.56	1.00
2	2.43	1.06	1.77	0.39	100.	0.226	15.7	7.85	25.21	8.28	0.19	0.49	0.32	0.97
3	6.31	9.88	1.91	0.25	11.9	0.108	20.0	7.41	25.92	42.9	0.87	0.10	0.03	1.57
4	3.49	2.81	1.66	0.50	100.	0.312	15.4	8.78	25.41	13.3	0.19	0.41	0.40	1.15
5	3.53	4.03	2.12	0.36	54.9	0.205	14.9	7.87	26.28	98.4	0.32	0.44	0.24	1.10
6	4.24	0.00	1.93	0.46	23.9	0.347	12.6	8.50	26.33	111	0.59	0.22	0.19	1.02
7	1.33	3.74	2.20*	0.29	8.37	0.137	40.3	7.00	26.53	175	0.26	0.53	0.21	1.20
8	3.12	7.35	1.82	0.15	24.1	1.380	3.44	7.09	25.85	36.6	0.58	0.35	0.06	1.39
9	1.30	1.36	1.71	0.71	12.9	0.360	11.1	6.96	25.94	45.0	0.84	0.05	0.11	17.2
10	1.74	0.00	1.91	0.32	100.	0.295	13.8	8.47	26.20	81.5	0.19	0.55	0.26	1.72
11	1.72	2.00	1.71	0.49	20.2	0.449	9.23	7.80	26.02	53.8	0.65	0.18	0.17	1.01
12	1.20	1.08	1.68	0.48	20.6	0.402	11.4	7.79	25.27	9.46	0.65	0.18	0.17	1.20
13	3.56	0.00	1.37	0.87	10.9	0.146	17.9	9.20	26.52	170	0.90	0.01	0.09	3.12
14	1.76	0.00	1.72	0.71	100.	0.294	16.0	8.24	25.82	33.6	0.19	0.23	0.58	1.07
15	1.31	2.43	2.20*	0.09	14.2	0.214	12.3	6.23	25.31	10.4	0.80	0.18	0.02	2.27
16	1.70	0.65	1.72	0.31	100.	0.327	13.0	8.33	25.85	36.2	0.19	0.56	0.25	1.44
17	0.65	0.85	1.74	0.14	48.7	0.326	11.4	7.97	25.85	36.5	0.35	0.56	0.09	1.08
18	1.45	0.19	2.20*	0.24	29.5	0.254	13.6	8.17	26.18	76.9	0.51	0.37	0.12	1.37
19	3.70	1.41	1.98	0.19	45.8	0.142	21.5	7.71	24.85	3.61	0.37	0.52	0.12	1.10
20	1.91	4.77	2.20*	0.36	9.63	0.210	16.8	6.46	25.36	11.9	0.94	0.04	0.02	1.39
21	1.77	0.00	1.79	0.75	22.7	0.206	19.6	7.98	26.09	63.4	0.61	0.10	0.29	3.59
22	2.75	8.84	1.86	0.21	50.5	0.108	25.1	7.84	25.42	13.5	0.34	0.52	0.14	1.09
23	1.59	0.00	1.41	0.45	86.9	0.626	9.59	7.99	25.02	5.40	0.22	0.43	0.35	0.99
24	1.63	0.00	1.82	0.94	32.2	0.182	32.2	8.26	25.96	46.5	0.48	0.03	0.49	2.13
25	2.34	0.00	1.79	0.40	25.7	0.351	12.9	8.49	26.26	94.2	0.56	0.27	0.17	1.83
26	2.31	7.25	2.10	0.03	33.8	0.310	9.69	7.37	26.23	87.7	0.46	0.52	0.02	1.14
27	2.75	0.00	1.85	0.22	37.6	0.554	8.29	7.50	25.44	14.2	0.43	0.45	0.12	1.12
28	1.45	0.00	1.69	0.60	71.3	0.353	13.7	8.24	25.81	33.0	0.26	0.30	0.45	1.26
29	1.18	1.36	2.00	0.12	30.9	0.389	8.85	7.76	26.03	55.1	0.49	0.45	0.06	1.24
30	1.87	2.64	2.20*	0.36	9.67	0.234	16.9	6.79	25.77	30.3	0.94	0.04	0.02	1.03
31	0.84	0.00	1.68	0.54	100.	0.404	12.9	8.70	25.84	35.9	0.19	0.37	0.43	0.99
32	0.90	0.14	1.80	0.44	100.	0.388	12.2	7.69	25.15	7.30	0.19	0.46	0.35	1.66
33	1.07	0.82	2.18	0.57	15.0	0.226	15.6	7.78	25.96	47.3	0.78	0.10	0.13	1.15
34	1.83	0.93	1.90	0.33	100.	0.252	14.8	8.71	26.03	55.1	0.19	0.54	0.26	1.11
35	1.76	0.90	1.80	0.83	100.	0.202	20.4	7.67	26.13	69.8	0.19	0.14	0.67	1.05
36	1.18	3.94	2.18	0.22	16.2	2.000	2.71	6.52	25.13	6.90	0.75	0.20	0.05	1.81
37	1.82	0.00	2.04	0.38	100.	0.219	17.2	8.23	25.88	39.3	0.19	0.50	0.31	1.33
38	1.42	0.37	1.58	0.97	100.	0.251	25.0	7.69	24.54	1.80	0.19	0.02	0.79	1.28
39	1.36	4.77	2.10	0.11	40.6	0.281	11.4	7.01	25.17	7.57	0.40	0.53	0.07	1.90
40	0.77	5.21	2.05	0.06	24.0	0.930	4.28	7.41	26.26	93.6	0.59	0.39	0.02	2.27
41	1.81	0.00	1.90	0.39	28.9	0.298	14.0	8.39	26.10	65.2	0.52	0.30	0.19	1.63
42	2.86	3.29	1.84	0.41	100.	0.083	31.3	7.74	24.68	2.45	0.19	0.47	0.33	1.01
43	2.69	0.00	1.71	0.58	55.8	0.406	11.9	8.07	26.10	64.7	0.31	0.29	0.40	1.29
44	2.78	5.90	2.17	0.04	27.6	0.501	6.71	7.26	26.13	68.6	0.53	0.45	0.02	2.33
45	1.46	0.00	1.82	0.49	41.0	0.286	14.1	8.62	26.75	290	0.40	0.30	0.30	2.42
46	4.02	0.55	1.81	0.81	100.	0.207	20.3	8.56	25.58	19.4	0.19	0.15	0.66	1.12
47	3.78	16.69	1.82	0.25	100.	0.115	29.8	7.96	25.62	21.5	0.19	0.61	0.20	0.99
48	2.11	0.87	1.85	0.19	18.1	0.525	8.61	7.19	25.16	7.40	0.70	0.24	0.06	1.19
49	4.90	0.36	2.20*	0.33	72.5	0.211	19.6	7.73	25.15	7.33	0.25	0.50	0.25	1.15
50	4.51	0.00	2.20*	0.80	7.88	0.131	48.5	7.86	27.42	1350	0.98	0.00	0.01	1.39
51	2.91	1.53	1.79	0.95	100.	0.112	45.2	7.65	25.32	10.8	0.19	0.04	0.77	1.38

Table 3. Broad-band SED key parameters. ID: object number, the same as Table 1; $\Gamma_{2-10\,\mathrm{keV}}$: the slope of the single power law fitted to the 2–10 keV spectrum. $L_{2-10\,\mathrm{keV}}$: 2–10 keV luminosity (in 10^{44} erg s^{-1}); $\kappa_{2-10\,\mathrm{keV}}$: the 2–10 keV bolometric correction coefficient; $\lambda L_{2500\,\text{Å}}$: the monochromatic luminosity at 2500 Å (in 10^{43} erg s^{-1}); $\nu L_{2\,\mathrm{keV}}$: the monochromatic luminosity at 2keV (in 10^{43} erg s^{-1}); α_{ox}: the optical X-ray spectral index; λL_{5100}: the monochromatic luminosity at 5100 Å (in 10^{44} erg s^{-1}); κ_{5100}: the 5100-Å bolometric correction coefficient; $\mathrm{FWHM}_{\mathrm{H}\beta}$: the narrow component subtracted Hβ FWHM; and $L_{\mathrm{bol}}/L_{\mathrm{Edd}}$: the Eddington ratio.

ID	$\Gamma_{2-10\,\mathrm{keV}}$	$L_{2-10\,\mathrm{keV}}$ ($\times10^{44}$)	$\kappa_{2-10\,\mathrm{keV}}$	$\lambda L_{2500\,\text{Å}}$ ($\times10^{43}$)	$\nu L_{2\,\mathrm{keV}}$ ($\times10^{43}$)	α_{ox}	λL_{5100} ($\times10^{44}$)	κ_{5100}	$\mathrm{FWHM}_{\mathrm{H}\beta}$ (km s^{-1})	$L_{\mathrm{bol}}/L_{\mathrm{Edd}}$
1	1.69 ± 0.06	4.941	11.9	81.3	25.6	1.19	8.15	7.24	13 000	0.11
2	1.67 ± 0.10	0.469	17.7	18.4	2.47	1.33	0.791	10.5	6220	0.089
3	1.77 ± 0.07	0.289	149	41.0	1.91	1.51	1.35	31.7	2310	1.3
4	1.80 ± 0.11	0.567	23.6	12.8	3.15	1.23	1.91	6.98	10 800	0.017
5	2.10 ± 0.22	2.284	43.2	134	12.9	1.39	5.48	18.0	2720	1.0
6	1.93 ± 0.18	4.855	22.9	290	27.6	1.39	14.8	7.52	5430	0.27
7	2.39 ± 0.22	0.267	657	61.3	2.43	1.54	1.95	89.6	1980	13
8	1.84 ± 0.04	0.418	87.7	23.5	2.89	1.35	0.539	68.1	2840	2.3
9	1.76 ± 0.01	0.839	53.8	22.7	5.35	1.24	0.113	399	3030	3.8
10	1.92 ± 0.05	3.532	23.1	205	23.1	1.36	7.59	10.8	4810	0.21
11	1.71 ± 0.11	1.811	29.8	78.9	9.03	1.36	3.75	14.4	5640	0.66
12	1.68 ± 0.23	0.502	18.9	21.2	1.57	1.43	1.04	9.12	4390	0.12
13	1.37 ± 0.12	0.751	227	790	2.99	1.93	42.6	4.00	10 800	0.082
14	1.69 ± 0.04	3.189	10.6	50.2	17.0	1.18	3.91	8.60	7060	0.15
15	2.35 ± 0.12	0.042	251	2.89	0.353	1.35	0.204	51.1	988	4.7
16	1.78 ± 0.07	1.502	24.2	90.8	8.24	1.40	4.26	8.53	3560	0.13
17	1.80 ± 0.20	0.779	46.9	71.7	3.62	1.50	3.31	11.1	2250	0.30
18	2.23 ± 0.08	1.254	61.5	157	9.67	1.46	6.11	12.6	2310	0.40
19	1.98 ± 0.18	0.084	43.1	8.59	0.497	1.47	0.443	8.19	2000	0.054
20	2.34 ± 0.12	0.053	224	4.44	0.476	1.37	0.215	55.4	774	3.1
21	1.70 ± 0.04	3.856	16.5	109	20.5	1.28	2.22	28.6	6090	0.51
22	1.70 ± 0.09	0.396	34.1	27.3	2.17	1.42	0.983	13.8	7050	0.15
23	1.80 ± 0.19	0.145	37.5	11.5	0.907	1.42	0.708	7.66	1980	0.043
24	1.83 ± 0.18	4.735	9.84	106	25.1	1.24	6.64	7.01	13 900	0.20
25	1.88 ± 0.03	3.054	30.9	249	20.0	1.42	8.44	11.2	4980	0.24
26	2.09 ± 0.25	0.362	243	63.3	2.60	1.53	2.04	43.2	1720	2.9
27	1.94 ± 0.04	0.277	51.5	20.3	2.51	1.35	0.988	14.4	4310	0.34
28	1.71 ± 0.14	2.951	11.2	63.6	13.2	1.26	4.80	6.91	4240	0.15
29	2.00 ± 0.12	0.726	76.0	76.3	4.75	1.46	3.25	17.0	3560	0.73
30	2.46 ± 0.09	0.146	207	13.4	1.28	1.39	0.452	67.2	954	3.8
31	1.69 ± 0.14	2.420	14.9	53.6	11.9	1.25	6.49	5.54	6810	0.055
32	1.88 ± 0.03	0.464	15.8	13.7	2.97	1.26	0.512	14.3	3100	0.12
33	2.14 ± 0.21	1.157	41.0	69.7	7.55	1.37	4.03	11.8	5690	0.60
34	1.90 ± 0.14	2.489	22.2	140	13.5	1.39	10.8	5.13	3310	0.082
35	1.76 ± 0.07	3.918	17.9	67.5	51.5	1.04	3.59	19.5	2790	1.2
36	2.20 ± 0.08	0.091	76.3	3.31	0.651	1.27	0.244	28.4	1890	1.6
37	1.95 ± 0.08	1.768	22.3	88.8	12.3	1.33	5.39	7.30	3960	0.18
38	1.55 ± 0.09	0.175	10.3	1.89	0.768	1.15	0.197	9.16	6630	0.028
39	2.17 ± 0.20	0.079	96.5	6.89	0.737	1.37	0.233	32.6	991	0.56
40	2.02 ± 0.06	0.468	200	70.0	3.54	1.50	2.05	45.7	2790	2.8
41	1.94 ± 0.05	2.444	26.7	167	15.8	1.39	6.26	10.4	2610	0.20
42	1.76 ± 0.11	0.158	15.5	4.92	0.804	1.30	0.265	9.28	4920	0.034
43	1.74 ± 0.07	4.524	14.3	109	25.7	1.24	4.36	14.9	4550	0.43
44	2.25 ± 0.05	0.236	292	45.5	2.13	1.51	2.36	29.2	1070	2.9
45	1.82 ± 0.06	17.502	16.6	645	98.4	1.31	30.4	9.58	10 900	0.53
46	1.81 ± 0.12	2.175	8.93	19.0	10.4	1.10	2.97	6.55	9930	0.041
47	1.45 ± 0.25	0.868	24.9	36.6	4.39	1.35	0.931	23.2	4100	0.18
48	2.03 ± 0.11	0.101	73.2	8.71	0.734	1.41	0.278	26.7	1190	0.37
49	2.40 ± 0.22	0.200	36.8	14.5	1.69	1.36	0.719	10.2	1340	0.11
50	2.41 ± 0.18	3.299	411	860	27.3	1.57	29.5	46.0	2200	14
51	1.67 ± 0.03	1.659	6.50	12.5	8.30	1.07	0.624	17.3	11 100	0.19

than 10 per cent of the peak effective transmission) are as follows: *UVW2*: 1805–2454 Å; *UVM2*: 1970–2675 Å; *UVW1*: 2410–3565 Å; *U*: 3030–3890 Å; *B*: 3815–4910 Å; *V*: 5020–5870 Å. We exclude the contribution from strong optical emission lines within the OM *U*, *B*, *V* bandpasses (and also the Balmer continuum contribution in the *U* band) by using the best-fitting optical underlying continuum which excludes such features from the SDSS spectral fitting. In fact, this was an important initial motivation of this study, that is, to obtain more accurate estimates of the true underlying continuum rather than simply to use the SDSS '*ugriz*' photometric

data. Inclusion of strong emission lines within these photometric data would result in overestimation of the optical continuum, and so compromise our aim to study the shape of the optical underlying continuum. This is an important spectral characteristic used to constrain the accretion disc component in the SED fitting (see also the discussion in Section 5.1.2). There are some strong emission lines within the UV bandpasses such as Lyα, C IV λ1549, C III λ1909 and Mg II λ2798, whose fluxes are not available from the SDSS spectrum. Accurate subtraction of these line fluxes for each object would require new UV spectroscopy. We conclude that inclusion of emission-line flux within the OM photometric points may account for some of the observed discrepancy.

(3) *Intrinsic source variability*. AGNs are well known to be variable across their SEDs. In general, there is a significant time-difference between acquisition of the SDSS and OM-UV data, so intrinsic variation may contribute to any observed discrepancy. Mrk 110 is the most extreme example of this phenomenon in our sample, as its SDSS spectrum has a very large discrepancy compared with the OM data. The recent paper by Landt et al. (2011) gives another set of optical spectra for Mrk 110, which is more consistent with our best-fitting model. It shows that the inclusion of OM data is helpful to identify cases such as this. As an additional test for variability, we assembled all available *GALEX* data for our sample. We find that 43 objects in our sample have *GALEX* data. Using a *GALEX* aperture of 12 arcsec, which is limited by the PSF and which is also similar to the UV-OM apertures, we compare these values with the SED model. The ratio of the *GALEX* data to our SED model within the same bandpass differs by less than a factor of 2 for the majority of our sample, and significantly the flux ratio distribution is almost symmetric and is centred close to unity. This suggests that the non-simultaneous OM and SDSS data are not likely to be a major impediment to our modelling.

In effect, these three factors will merge together to produce the observed discrepancy between the SDSS and OM data. Since the combined effects of points (1) and (2) will add flux and generally be greater than that caused by optical/UV variability as shown by previous long-term reverberation mapping studies (Giveon et al. 1999; Kaspi et al. 2000), we should treat the OM points included in our SED modelling as upper limits when interpreting the results of our modelling. Indeed, the 90 per cent confidence uncertainties in the BH masses derived directly from the XSPEC fitting are almost certainly small compared with the systematic errors introduced by the above uncertainties.

5.3.2 *The observed flat optical continuum*

A related problem in our fitting is about the SDSS continuum shape. For some AGNs, their SDSS continuum data points exhibit a very different spectral slope from that of the SED model. This cannot be reconciled by adjusting the parameters of the accretion disc model, and thus implies the presence of an additional component at longer optical wavelengths, which flattens compared with that predicted by the accretion disc models. One obvious explanation for this flux excess is the contribution from the host galaxy. In late-type host galaxies, such as elliptical and S0 galaxies, emission from their old stellar populations peaks at near-IR wavelengths. Kinney et al. (1996) combined spectra of quiescent galaxies and constructed an average spectral template for each morphological type, including bulge, elliptical, S0, Sa, Sb, Sc and starburst galaxies. For some objects in our sample with high-S/N SDSS spectra which show at least marginal stellar absorption features, we have added the corresponding type of host galaxy spectral template taken from Kinney et al. (1996) into the overall SDSS spectral fitting. This revised underlying continuum in the optical was then used in the broad-band SED fitting. We are then able to compare it with the original fit, to see how the subtraction of a stellar population template effects the overall SED fitting.

Fig. 5 shows two examples. The first is 2XMM J112328.0+052823, in which, after subtracting the host galaxy component, the observed optical continuum is closer to the slope of the SED model. However, the results for PG 1415+451 in Figs 5(c) and (d) imply that its host galaxy cannot be the origin of the flat optical spectrum. The reason is that its optical spectrum does not show any strong stellar absorption features. This means that the maximum amount of host galaxy contribution is small, and so there remains a substantial inconsistency in the slope versus the SED model. In addition to 2XMM J112328.0+052823 above, only Mrk 1018 and 2XMM J125553.0+272405 show clear stellar absorption features. Also, the 3-arcsec-diameter fibre excludes much of the host galaxy component at these redshifts. Therefore, on these general grounds, we conclude that host galaxy contamination is small for most sources in our sample, and consequently cannot fully account for the observed flat optical continuum. Additional support for this view comes from good correlations between the X-ray components and the red optical continuum, suggesting that this extra optical flux is likely related to the intrinsic activity (e.g. Soria & Puchnarewicz 2002; Collin & Kawaguchi 2004; Hao et al. 2010; Landt et al. 2011).

6 STATISTICAL PROPERTIES OF THE SAMPLE

Histograms of data on our sample are shown in Figs 6, 8 and 9, including redshift, H I column density, optical and X-ray modelling parameters, etc. The red region in the histograms shows the distribution for the 12 NLS1s in our sample. It is clear that NLS1s are distinct in the whole sample in several respects.

6.1 General properties

Fig. 6 shows some basic properties of our sample which are not model-dependent:

(1) *Redshift*. The sample's redshift ranges from 0.031 (Mrk 493) to 0.377 (HS 0810+5157). The NLS1s are found mainly at lower redshifts, with $\langle z \rangle_n = 0.12$ compared to $\langle z \rangle_n = 0.19$ for the BLS1s. For comparison, we see that the sample of VF07 has a similar redshift range, but it has a lower average redshift of 0.10.

(2) *The Galactic* $n_{\rm H}$. The average Galactic $n_{\rm H}$ is 2.25×10^{20}.

(3) *The photon indexes obtained from simple power-law fits to the restricted energy range of 2–10 keV*. The NLS1s cluster on the higher photon index side, with an average of 2.21 ± 0.20, which differs from the sample average of 1.92 ± 0.25 and the BLS1s' average of 1.83 ± 0.18. This means that NLS1s tend to have softer X-ray spectra, which is further confirmed in the following section on the mean SEDs.

(4) *The X-ray continuum and 2–10 keV luminosity*. This distribution shows that NLS1s have lower 2–10 keV luminosities in spite of their steeper slopes. We note that the VF07 sample has a similar distribution, except for their inclusion of three extremely low X-ray luminosity AGNs (i.e. NGC 4395, 3227 and 6814). These objects were not included in our sample due to our selection criteria and/or a lack of SDSS spectra.

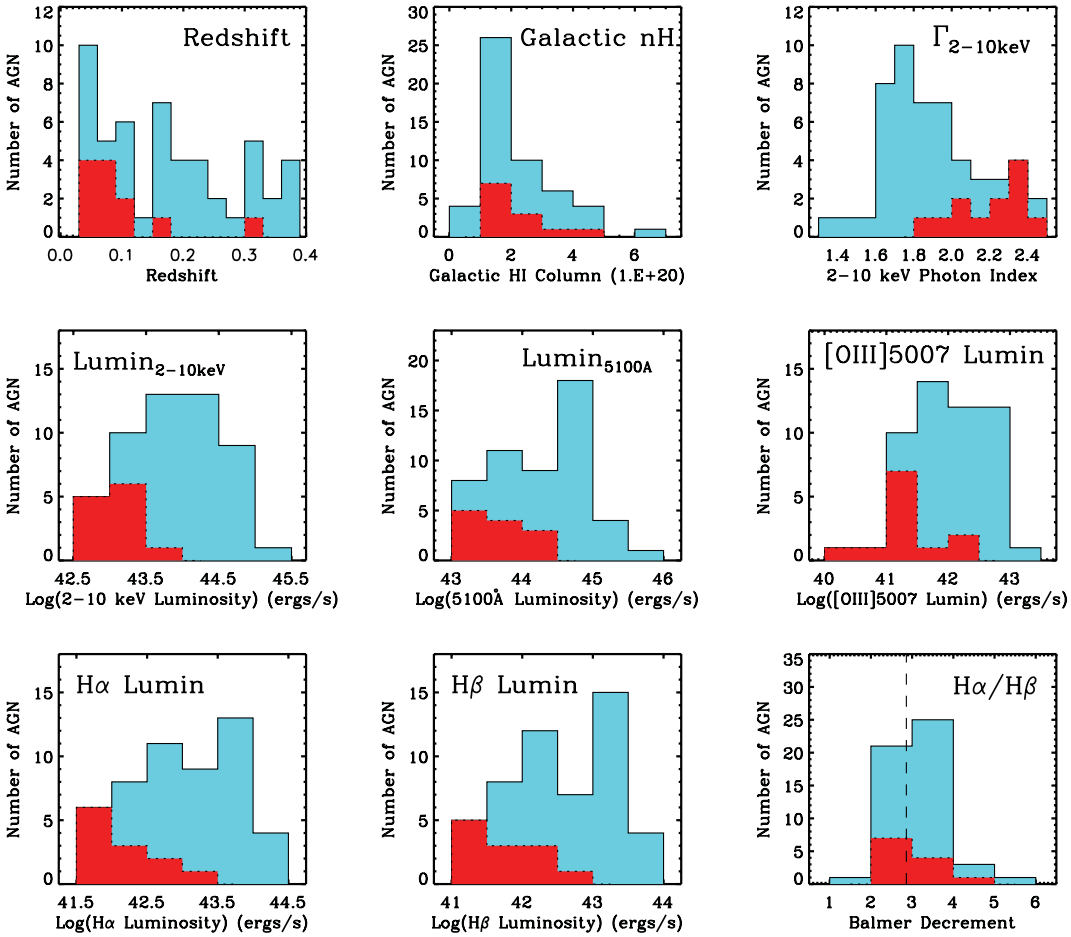

Figure 6. Distributions of our sample for different properties. In each panel, the blue areas show the distribution for the whole sample, while the red areas show the distribution for the 12 NLS1s in our sample. We note that the Hα, Hβ and [O III] λ5007 luminosities are based on results of line profile fitting, after subtracting the blends from other nearby emission lines (see Section 3.1). For comparison, we also indicate the Balmer decrement value of 2.86, found under case B recombination, as shown by the dashed line in the bottom rightmost panel.

(5) *The optical continuum luminosity at 5100 Å*. On average, the NLS1s have lower optical luminosities than BLS1s.

(6–8) *The [O III] λ5007, Hα and Hβ emission-line luminosities*. Again the NLS1s have on average lower luminosities than BLS1s.

(9) *The Balmer decrement*. The average value for the whole sample is 3.14 ± 0.62, and for NLS1s it is 3.05 ± 0.38. This difference is not statistically significant, but we will return to the issue in our next paper (Paper II), where we consider the separate components as well as the overall profile.

6.2 Results from the broad-band SED modelling

Fig. 7 shows the following properties derived from the SED fits:

(1) *The bolometric luminosity*. The distribution range is between 1.8×10^{44} (Mrk 464) and 1.4×10^{47} erg s^{-1} (PG 2233+134). There is no clear difference in the distribution of the complete sample and the subset of NLS1s. The average luminosity is $\log(L_{\rm bol}) = 45.49 \pm 0.55$, which is consistent with the value of 45.19 ± 1.01 found in the VF07 sample, except for the three extremely nearby and low-luminosity AGNs in VF07.

(2) *The black hole mass*. Using the best-fitting black hole masses, the whole sample peaks between 10^7 and 10^8 M$_\odot$. Equation (5) suggests that the black hole mass should depend on both Hβ FWHM and L_{5100}, and the results from our SED fitting suggest that NLS1s with smaller Balmer line FWHM do indeed harbour lower mass black holes. KUG 1034+396 has the lowest black hole mass in our sample. The value of 1.7×10^6 M$_\odot$ is consistent with the estimate based on the first firmly detected AGN quasi-periodic oscillation found in this source (Gierliński et al. 2008). Again, we can compare our results with those of the VF07 sample. We find that their average black hole mass is 7.89 ± 0.82, calculated using the $M(L_{5100}, {\rm FWHM}_{{\rm H}\beta})$ relation. Adopting the same method for our sample, we find a very similar average of 7.99 ± 0.93. Our best-fitting masses have a slightly lower average value of 7.83 ± 0.64 (also see Section 6.5 for a comparison of different estimates of black hole masses).

(3) *The Eddington ratio*. The average values are 3.21 ± 3.07 for NLS1s which display a wide dispersion, 0.57 ± 0.50 for BLS1s and 0.93 ± 0.85 for the whole sample. Of the eight objects whose Eddington ratios are above 1, six are NLS1s, and the highest value is 14.2 (PG 2233+134). Clearly, NLS1s tend to have higher Eddington ratios. Our Eddington ratio distribution is also similar to that found in the sample of VF07 whose average value is 0.47 ± 0.44, except that their distribution has a more pronounced peak at $\sim$0.1.

(4) *The α_{ox} index*. It is defined between rest-frame continuum points at 2500 Å and 2 keV (see Lusso et al. 2010, and references

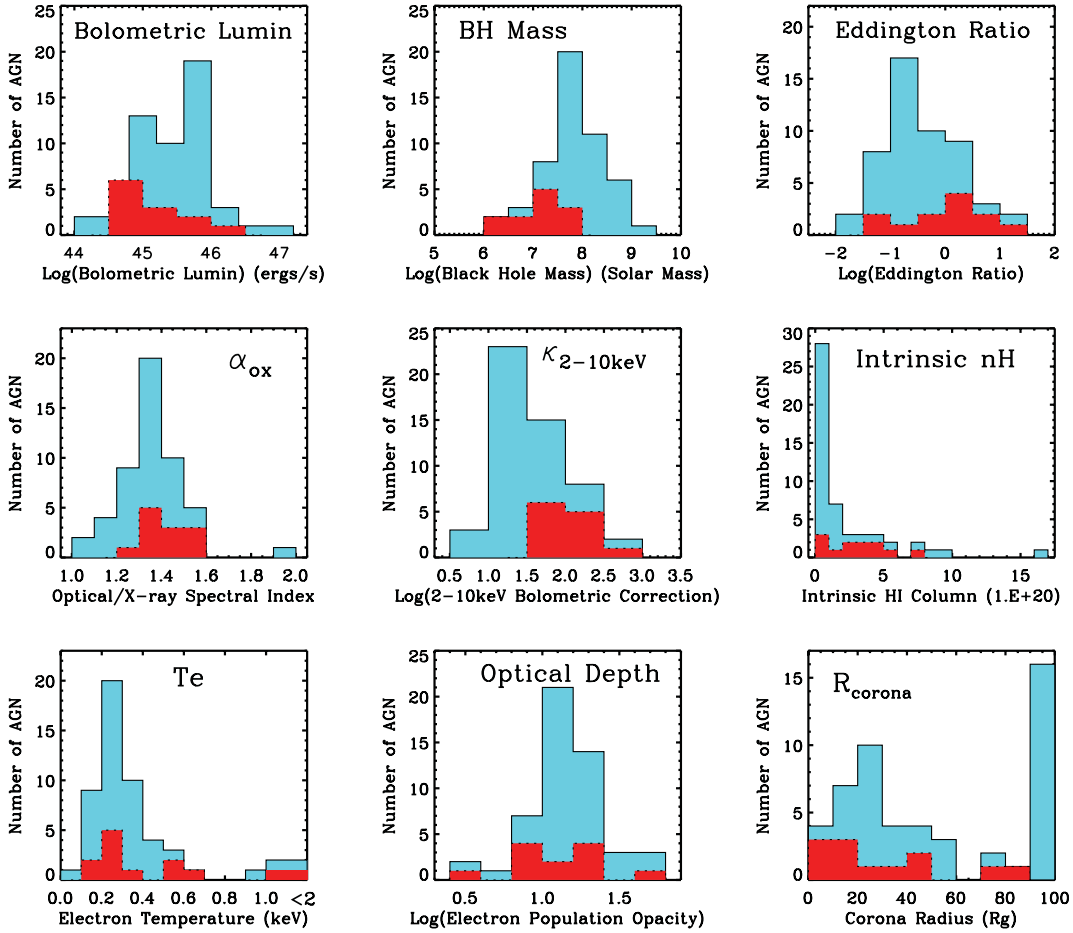

Figure 7. The distribution of model-dependent parameters using the same colour coding as in Fig. 6. Comments on each distribution are given in Section 6.2.

therein). The distribution for NLS1s is peaked at marginally higher values than for BLS1s.

(5) *The κ_{2-10} bolometric correction*. It is defined as L_{bol}/L_{2-10} (see VF07, and references therein). We find that NLS1s have a significantly higher fraction of their bolometric luminosity emitted as hard X-rays than the BLS1s. Compared with the VF07 sample, both distributions peak at $\kappa_{2-10} = 10$–30, but our sample shows a smoother distribution decreasing as κ_{2-10} increases after ∼30, and so results in a slightly higher average value of κ_{2-10}.

(6) *The intrinsic n_H*. This distribution shows that the intrinsic equivalent neutral hydrogen column densities are low for our sample, which is a natural consequence of our initial sample selection criteria. The NLS1s have slightly higher intrinsic absorption than BLS1s, which may imply a slightly higher dust reddening. However, the distribution of Balmer decrements shows no significant difference between these two types of AGNs.

(7) *The temperature of the Comptonization component used to describe the soft X-ray excess*. This is close to 0.2 keV in all objects, confirming the trend seen in previous studies for this component to exhibit a narrow range of peak energy (Czerny et al. 2003; Gierliński & Done 2004). The distribution peak at this energy is more marked for the BLS1s than for NLS1s, although the small number statistics mean that this difference cannot be considered as definitive for our sample.

(8) *The optical depth of the soft-excess Comptonized component*. It is clear that this component is always optically thick, with most objects having $\tau \sim 10$–30. There is no significant difference in temperature or optical depth between the broad- and narrow-line objects.

(9) *The difference in corona radius distribution between the BLS1s and NLS1s*. The corona radius controls the relative amount of power emerging from the accretion disc and the soft X-ray excess/hard tail. There are two peaks in the distribution for the broad-line objects, one between 10 and $20R_g$ (where $R_g = GM/c^2$), and the other at $100R_g$ (which is set as the upper limit of this parameter in our broad-band SED model). By contrast, these radii in NLS1s are consistent with just the first peak. At first sight, this is surprising, since NLS1s are expected to be those with the strongest soft X-ray excess. However, their similar soft-excess temperatures around 0.2 keV suggest that atomic processes may be significant (reflection and/or absorption from partially ionized material), and this may influence our fits. The average corona radii are $(32 \pm 26)R_g$ for NLS1s, $(59 \pm 37)R_g$ for BLS1s and $(53 \pm 36)R_g$ for the whole sample. This supports the conclusion of VF07 that high-Eddington-ratio AGNs have lower corona fractions compared to those with low Eddington ratios.

6.3 Balmer line parameter distribution

Fig. 8 shows further details of the modelled profiles of Hα (first row) and Hβ (second row).

(1) *The FWHM of the broad emission profile*. This is calculated by co-adding the two best-fitting Gaussian profiles for the broad- and intermediate-line components, and then using the resultant profile to

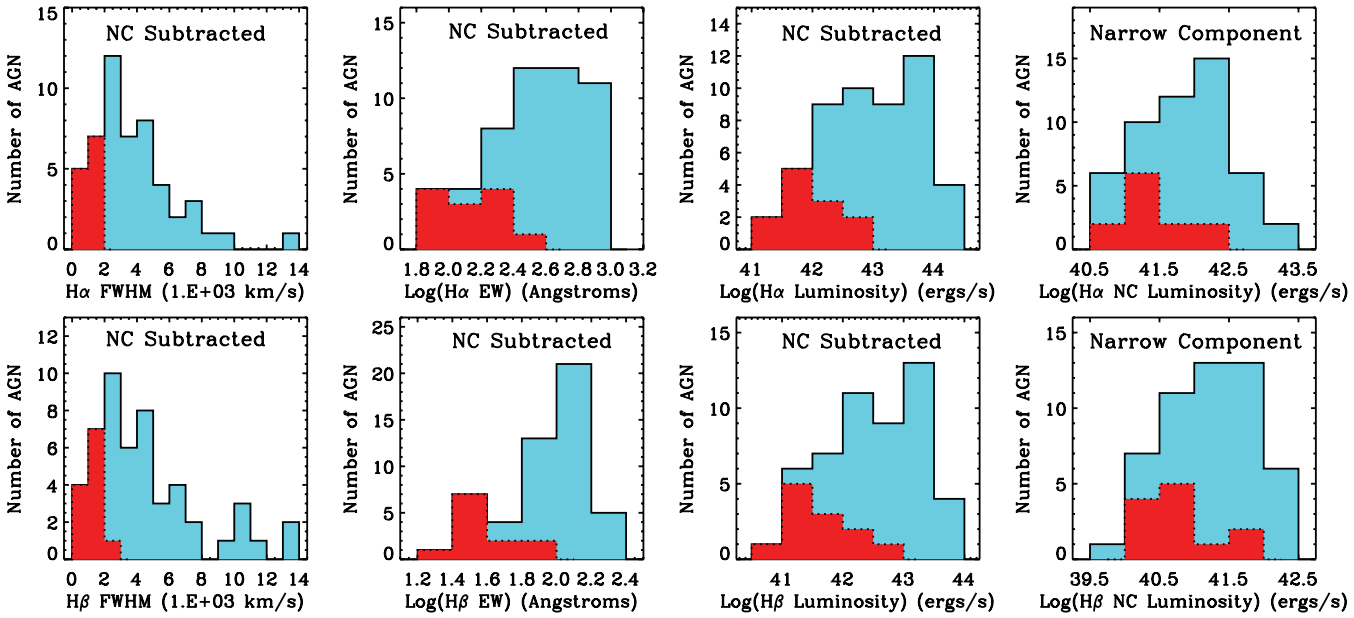

Figure 8. The Balmer line parameter distributions. The first row is for Hα and the second is for Hβ. We combine the intermediate and broad components in each Balmer line profile to form the total broad-line properties, giving values of the FWHM, EW and luminosity. The panels in the rightmost column show the luminosity distribution of the narrow component for comparison. The distributions for the 12 NLS1s are indicated by the red regions, as in Fig. 6.

determine the FWHM. This is equivalent to subtracting the narrow-line core from the observed profile and measuring the resultant FWHM.

Note, the NLS1s by definition have Hα $< 2000\,\mathrm{km\,s^{-1}}$.

(2–3) The EWs and line luminosities are again measured using the total broad emission-line profile as above. The NLS1s have both lower EWs and line luminosities.

(4) By contrast, there is no pronounced difference between NLS1s and BLS1s in their Balmer narrow-line component. This suggests that the NLR is less influenced by whatever difference in properties is responsible for the defining difference between NLS1s and BLS1s in the BLR.

6.4 The bolometric luminosities

The fraction of the total luminosity contained in each component of the SED model is shown in Fig. 9. The upper left-hand panel shows these fractions as a function of bolometric luminosity. It seems that as the bolometric luminosity increases, the disc component slightly increases in importance. However, the total number of objects at high luminosities is small, as seen in the upper right-hand panel, where the fraction is multiplied by the number of objects in the bin, so we should be cautious about this finding.

The lower panel shows this fraction for each of the objects ranking from the smallest to biggest Hβ FWHM. Thus, low-rank objects have the narrowest Hβ (and hence are by definition NLS1s). These also have the lowest black hole masses and highest Eddington ratios. They are more likely to have a smaller fraction of their total luminosity emitted in the soft-X-ray-excess component than the BLS1s. This relates to the issue of the corona radii [see point (9) of Section 6.2]. There are also some BLS1s which have an apparently high fraction of power in their soft X-ray excesses, but they may also have alternative spectral fits including reflection and/or absorption.

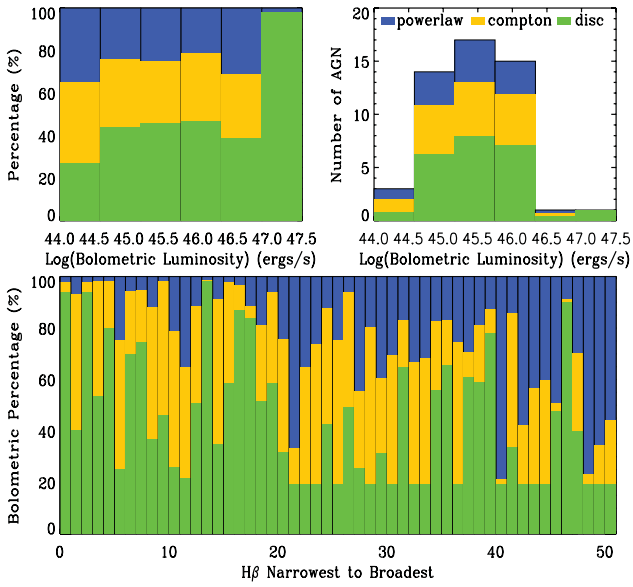

Figure 9. The bolometric luminosity distribution for the different continuum components of the SED, that is, accretion disc (green), Comptonization (orange) and hard X-ray Comptonization (blue). The upper left-hand panel shows the percentage within each luminosity bin for each of these three SED components. The upper right-hand panel shows the luminosity distribution of the whole sample, with each bin divided into three regions according to the fractional contribution from the different components in that luminosity bin. The lower panel shows how the contribution from each component changes as a function of rank order in Hβ FWHM, after the narrow-line component has been removed.

We note that in all these plots the lower limit to the disc fraction of 0.19 results from setting an upper limit of $100R_{\mathrm{g}}$ for the corona radius parameter, as mentioned in Section 5.2.

6.5 The black hole mass

The black hole mass is one of the key parameters used in our SED fitting, and it largely determines the continuum shape in the optical/UV region. The masses derived from reverberation mapping are considered to be the most accurate, but the total number of objects which have been studied using this technique is still relatively small (e.g. Peterson et al. 2004; Bentz et al. 2010; Denney et al. 2010). In the absence of reverberation mapping, the empirical relation between $M_{\rm BH}$ and Hβ linewidth and L_{5100} is often used as a proxy to estimate the black hole mass (Peterson et al. 2004). A serious limitation of this method is that it is still not clear which specific measure of the Hβ profile provides the closest association with the velocity dispersion of the gas in the BLR.

There are various alternative measures of the velocity width used for determining the black hole mass, including the FWHMs of the intermediate component (IC) and the broad component (BC) (e.g. Zhu et al. 2009). One could also use the model-independent second momentum (e.g. Peterson et al. 2004; Bian et al. 2008), or more simply the FWHM of the Hβ line after subtracting the narrow component (NC) (e.g. Peterson et al. 2004). The NC-subtracted FWHM and the second momentum estimates often lie within the range of values covered by the IC and BC FWHMs, except for some peculiar objects such as those with broad double-peaked profiles, for example, UM 269. Given all these uncertainties, we decided to adopt the best-fitting black hole mass obtained from the SED model, rather than simply fixing it at a value determined from a specific linewidth measurement. Moreover, it is now suggested that radiation pressure may be important in modifying the black hole mass derived using the relation between $M_{\rm BH}$ and L_{5100} and Hβ FWHM, especially for objects with high Eddington ratios such as most NLS1s (e.g. Marconi et al. 2008).

In order to compare our results with those from other studies, we have made various estimates of black hole mass for every source in our sample as follows:

(1) $M_{\rm BH,IC}$, $M_{\rm BH,BC}$ and $M_{\rm BH,IC+BC}$ are derived using equation (5) with different Hβ FWHMs obtained from our Balmer-line-fitting procedure.

(2) $M_{\rm BH,\sigma}$ is the black hole mass calculated from the second momentum of the total Hβ line profile (see Peterson et al. 2004 for details of the definition of 'second momentum'), by using $R_{\rm BLR} \propto L_{5100}^{0.518}$ and a geometry factor of $f = 3.85$. These assumptions are considered to be appropriate when using second momentum as a measure of the velocity dispersion in the BLR (Bentz et al. 2006; Collin et al. 2006; Bian et al. 2008).

(3) $M_{\rm BH,RP}$ is the black hole mass corrected for radiation pressure, using equation (9) in Marconi et al. (2008) with $f = 3.1$, $\log(g) = 7.6$.

We compare the black hole mass distributions obtained from these different methods in Fig. 10. The mean values are listed in Table 4.

$M_{\rm BH,IC}$ and $M_{\rm BH,BC}$ represent the two extreme estimates of black hole mass. $M_{\rm BH,IC}$ could still be influenced by contamination from a NLR component, especially for NLS1s where deconvolution of the narrow and broad components is very difficult. If there is a residual narrow-line component, it will introduce a bias that underestimates black hole masses. Conversely, using $M_{\rm BH,BC}$ is more likely to bias towards higher black hole masses, due to the presence of low-contrast very broad wings often seen in Hβ profiles. We found $\rm FWHM_{IC+BC}/\sigma_{H\beta} = 1.30 \pm 0.39$ for our sample, which is consistent with 1.33 ± 0.36 found by Bian et al. (2008). This leads to slightly lower values of $M_{\rm BH,\sigma}$ than $M_{\rm BH,IC+BC}$, but these two methods both give black hole masses between $M_{\rm BH,IC}$ and $M_{\rm BH,BC}$, with $M_{\rm BH,IC+BC}$ spanning a broader mass range.

Our best-fitting SED black hole masses ($M_{\rm BH,FIT}$) are also distributed between $M_{\rm BH,IC}$ and $M_{\rm BH,BC}$, with similar average masses to $M_{\rm BH,IC+BC}$ (a comparison is shown in Fig. 11a). Note that $M_{\rm BH,FIT}$ is a free parameter in the SED fitting unless it hits the lower or upper limits set by $M_{\rm BH,IC}$ and $M_{\rm BH,BC}$, which occasionally happened (see Table C1). It is clearly shown in Fig. 11 that the black hole masses

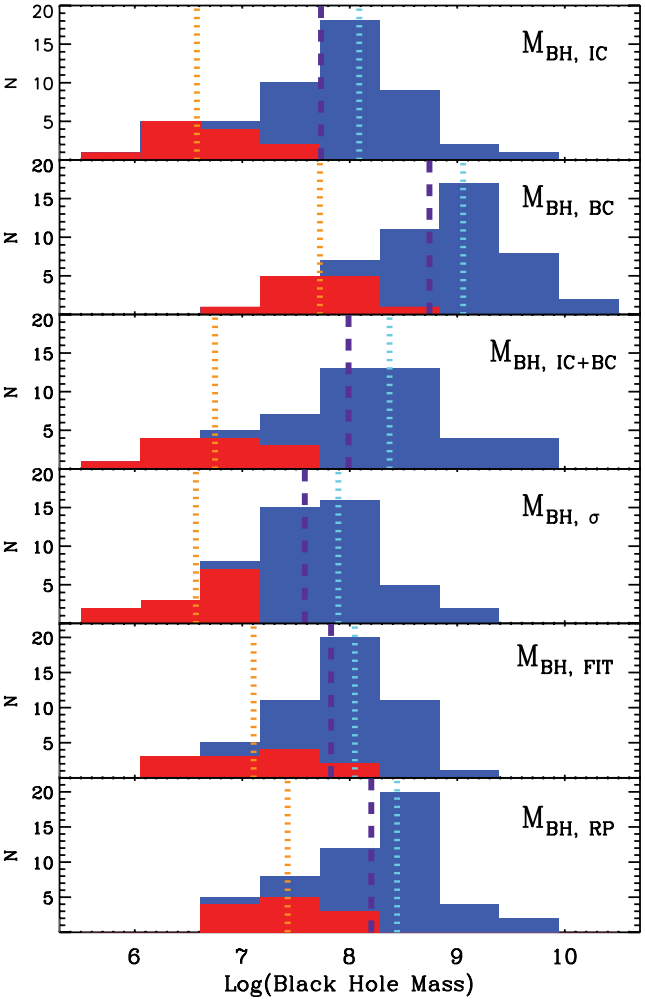

Figure 10. A comparison of various methods used to derive black hole mass. The total distributions are shown with the 12 NLS1s shown by the red regions. The purple dashed line indicates the average black hole mass for the whole sample. The orange and cyan dotted lines indicate the average masses of NLS1s and BLS1s, respectively. The average values are listed in Table 4. Values for individual objects are listed in Table C1 (Supporting Information with the online version of this paper).

Table 4. The average black hole masses, as shown in Fig. 10.

	NLS1	BLS1	All
$\langle M_{\rm BH,IC}\rangle$	6.58 ± 0.49	8.09 ± 0.56	7.73 ± 0.84
$\langle M_{\rm BH,BC}\rangle$	7.72 ± 0.49	9.05 ± 0.55	8.74 ± 0.78
$\langle M_{\rm BH,IC+BC}\rangle$	6.75 ± 0.49	8.37 ± 0.65	7.99 ± 0.93
$\langle M_{\rm BH,\sigma}\rangle$	6.57 ± 0.46	7.89 ± 0.47	7.58 ± 0.73
$\langle M_{\rm BH,FIT}\rangle$	7.11 ± 0.54	8.05 ± 0.48	7.83 ± 0.64
$\langle M_{\rm BH,RP}\rangle$	7.42 ± 0.39	8.44 ± 0.53	8.20 ± 0.66

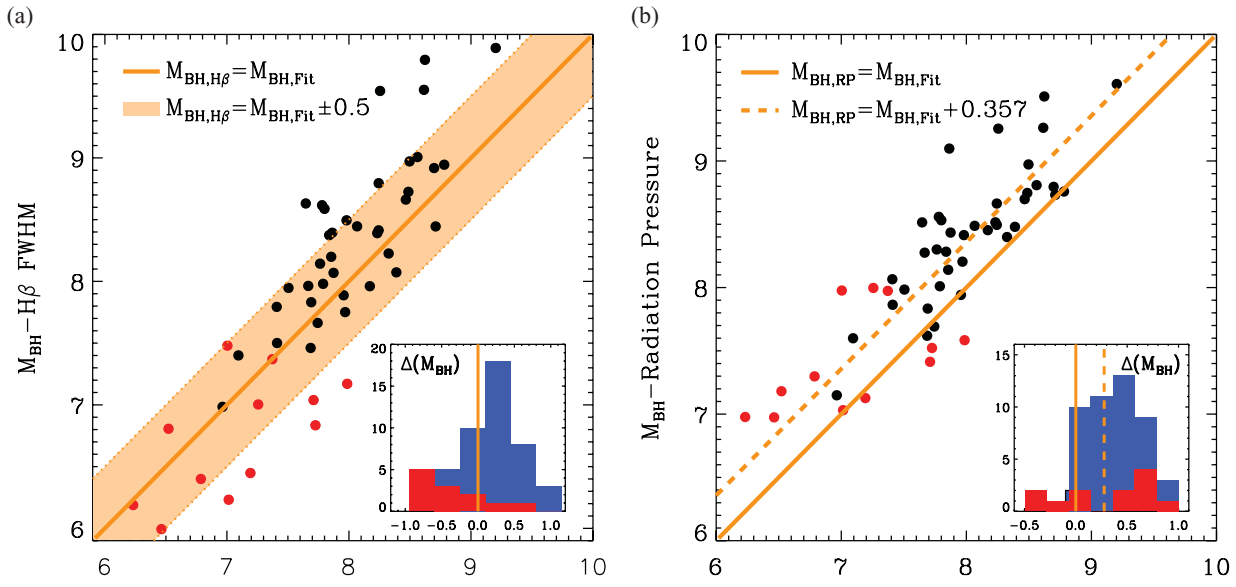

Figure 11. Correlations between best-fitting black hole mass ('M_{BH}-fitting' or '$M_{BH,FIT}$'), Hβ FWHM determined black hole mass ('M_{BH}-Hβ FWHM' or '$M_{BH,IC+BC}$') and radiation pressure corrected black hole mass ('M_{BH}-radiation pressure' or '$M_{BH,RP}$'). The red points represent the 12 NLS1s. The inset in panel (a) shows the distribution of the mass difference between $M_{BH,IC+BC}$ and $M_{BH,FIT}$, while the inset in panel (b) shows the distribution of the mass difference between $M_{BH,RP}$ and $M_{BH,FIT}$. The red regions highlight the distribution of NLS1s.

from the SED fitting are not consistent with estimates based on either extremely narrow or extremely broad lines. Thus, for NLS1s, the mean $M_{BH,FIT}$ is 0.36 dex higher than $M_{BH,IC+BC}$, while for BLS1s, the mean $M_{BH,FIT}$ is 0.22 dex lower than $M_{BH,IC+BC}$. Interestingly, this also implies that $M_{BH,FIT}$ of NLS1s may have less deviation from the established M–σ_* relation than that using the $M(L_{5100}, \mathrm{FWHM}_{H\beta})$ relation as shown in several previous studies (e.g. Wang & Lu 2001; Bian & Zhao 2004; Zhou et al. 2006).

The situation may be further complicated as Marconi et al. (2008) showed that NLS1s could be consistent with the M–σ_* relation if a correction for radiation pressure is applied to black hole masses derived from the $M(L_{5100}, \mathrm{FWHM}_{H\beta})$ relation. In our sample, correction for radiation pressure adds to the average $M_{BH,IC+BC}$ by 0.67 dex for NLS1s, 0.07 dex for BLS1s and 0.21 dex for the whole sample. We also found a very similar mass distribution between $M_{BH,RP}$ and $M_{BH,FIT}$, except for an average of 0.36 dex higher in $M_{BH,RP}$. The differences between the average masses of NLS1s and BLS1s are 0.78 and 0.72 dex in $M_{BH,RP}$ and $M_{BH,FIT}$, respectively (see Fig. 11b). Therefore, if $M_{BH,RP}$ can provide a good match to the M–σ_* relation even down to low-mass NLS1s as proposed by Marconi et al. (2008), then our SED-determined $M_{BH,FIT}$ may also give similar results. This implies that the suggested deviation from the M–σ_* relation for NLS1s may not be an intrinsic property, but rather a consequence of using black hole estimates based on the $M(L_{5100}, \mathrm{FWHM}_{H\beta})$ relation, which may not be appropriate for NLS1s (e.g. Grupe & Mathur 2004; Komossa 2008).

6.6 The average SEDs

Elvis et al. (1994) constructed SED templates for both radio-loud and radio-quiet AGNs, based on a sample of 47 quasars between redshift 0.025 and 0.94. VF07 modelled optical-to-X-ray SEDs for a sample of 54 AGNs with redshifts between 0.001 and 0.371, and showed that the SED is related to Eddington ratio. They also suggested that $\kappa_{2-10\,\mathrm{keV}}$ is well correlated with Eddington ratio. In a later study (Vasudevan & Fabian 2009) based on SED modelling of 29 local AGNs from Peterson et al. (2004), the SED dependence on Eddington ratio was reinforced. Recently, Lusso et al. (2010) studied 545 X-ray-selected type 1 AGNs over the redshift range of 0.04–4.25. They computed SEDs at different redshifts, and investigated α_{ox} correlations with other parameters such as redshift, $\kappa_{2-10\,\mathrm{keV}}$, λ_{Edd}, etc.

We present a mean SED for our sample which is subdivided according to their Hβ FWHM. This gave three subsamples, those with the narrowest lines, those with moderately broad lines and those with very broad lines. All objects were de-redshifted to their local frame. First, each of the best-fitting SEDs was divided into 450 energy bins between 1 eV and 100 keV. For each energy bin, we calculated the monochromatic luminosity for the subsample with 12 NLS1s, using their individual SED models. Then, an average value and standard deviation in each energy bin were calculated in logarithm space. Thus, a mean SED for the 12 NLS1s was constructed. Using the same method for the 12 moderate and 12 broadest line objects, their mean SEDs were produced. The total SED energy range is 1 eV to 100 keV, but we note that only spectral ranges from 1.5 to 6 eV and 0.3 to 10 keV are actually covered by the observational data, and all other ranges are based on model extrapolations.

Obviously, limitations of our mean SEDs include the relatively small sample sizes comprising the SEDs, and the redshift restriction $z < 0.4$. On the other hand, we have assembled high-quality data sets of optical, UV and X-ray observations. The exclusion of objects with high intrinsic absorption in the optical/UV helps to simplify the modelling assumptions. Our exclusion of warm-absorber objects may have introduced unknown selection effects, but again this simplified the SED modelling. Our model of the accretion flow also includes more detailed physical assumptions on the optical-to-X-ray spectrum than in previous broad-band SED studies. These advantages make our broad-band SED fitting more physically plausible. Thus, our mean SEDs too should be more reliable, especially in the unobservable far-UV region, where often the peak of the energy is emitted.

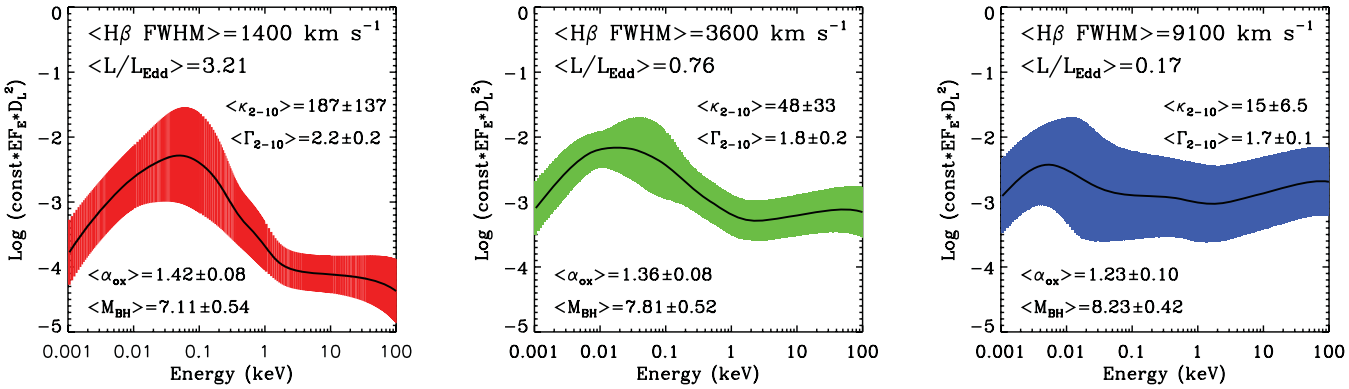

Figure 12. The average SED of our sample. The left-hand panel shows the averaged SED for the 12 NLS1s (including two marginal NLS1s, 2XMM 112328.0+052823 and 1E 1346+26.7). The average Hβ FWHM is 1400 $\pm$ 500 km s^{-1}. The red area indicates a 1 standard deviation region on either side of the average spectrum. The central panel is for 12 objects with moderate linewidth. The average FWHM is 3700 $\pm$ 600 km s^{-1}. The green region indicates 1 standard deviation. The right-hand panel shows the mean SED for the 12 broadest line objects in our sample, including the one double-peak source. The average FWHM is 9800 $\pm$ 2900 km s^{-1}. We also show the average value of the 2–10 keV power-law photon index, the 2–10 keV bolometric correction, and the α_{ox} value with 1σ error. D_L in the Y-axis title is the luminosity distance. The unit of the Y-axis is 'keV (erg s^{-1} keV^{-1})' in logarithm. The same arbitrary constant of 1.31 $\times$ 10^{-46} is used for rescaling each plot.

Fig. 12 shows the mean SEDs for the three subsets of our sample. We caution that there is still substantial spectral diversity within each subsample, and echoes Elvis et al.'s (1994) warning that if AGN SEDs are simply averaged without considering their detailed intrinsic properties, then the dispersion in the resultant mean SED will be large, so the mean SED may lose some useful information about AGN properties. Nevertheless, there appears to be a clear SED connection with Hβ FWHM. As the linewidth increases, so the BBB in the UV region becomes weaker relative to the hard X-rays, and its peak shifts towards lower energy. Also the spectral slope at high energies becomes harder.

This evolution in spectral shape is similar to that found by VF07 and Vasudevan & Fabian (2009), in which two mean SEDs of different mean Eddington ratios were compared. This relation might be expected since the FWHM and Eddington ratio are also strongly (anti)correlated in our sample. VF07 interpreted the spectral diversity as a scaled-up version of the different accretion states of Galactic black hole binaries. The low-Eddington-ratio AGN could be analogous to the low/hard state in black hole binaries in having a weak disc giving a strong high-energy tail, and the high-Eddington-ratio sources are analogous to the high/soft state, in which the disc emission dominates. Our SED templates do not extend down to such low Eddington ratios as in VF07, but we still see a similar behaviour.

7 SUMMARY AND CONCLUSIONS

In this paper, we presented a spectral study of 51 unobscured type 1 AGNs, including 12 NLS1s. We assembled X-ray data from the EPIC monitor onboard the *XMM–Newton* satellite, and optical data from the SDSS DR7. In addition, we added optical/UV data from the *XMM–Newton* OM monitor when available. Our results confirm some previously known correlations. For example, NLS1s often have softer power-law fits from 2 to 10 keV, and have lower 2–10 keV luminosities. Their Hα, Hβ and [O III] λ5007 lines are also less luminous on average than found in BLS1s.

We use detailed models to fit the Hα and Hβ line profiles, with multicomponents to deblend the narrow, intermediate and broad components by means of simultaneous modelling of the Fe II continuum and other blended lines. We then use results from the Hβ line fitting to constrain the black hole mass. The FWHM of the intermediate and broad components gives a lower and upper limit for the mass, respectively. This supports previous studies which find that NLS1s tend to have lower black hole masses and higher Eddington ratios, although their bolometric luminosities are not significantly different from those of BLS1s.

We include the Balmer continuum and permitted iron emission, and extend the modelling across the entire SDSS spectrum in order to isolate the intrinsic optical underlying continuum. However, this pure optical continuum is often (in 32/51 objects) flatter than is predicted by the standard accretion disc model. This could indicate some contamination from the host galaxy, but the lack of stellar absorption features in most of the SDSS spectra suggests that this cannot be a general explanation. Instead, it seems more likely that there is an additional component in the optical region related to the AGN, which is not yet well understood.

We also show that the Balmer continuum is not well modelled if the edge wavelength is fixed at its laboratory value of 3646 Å. It is shifted redwards and smoothed by more than predicted by the FWHM of the Balmer emission lines. These effects could both be produced by density broadening. Potentially more detailed models of the optical emission could employ this as a new diagnostic tool for studying the physical conditions, for example, electron density and temperature, in the innermost Balmer emitting regions.

The optical, UV and X-ray data were fitted using a new broad-band SED model, which assumes that the gravitational potential energy is emitted as optically-thick blackbody emission at each radius down to some specific corona radius. Below this radius, the remaining energy down to the last stable orbit is divided between a soft-X-ray-excess component and a hard X-ray tail. This energetically constrains the model fits in the unobservable extreme-UV region. We construct the resulting SEDs for each of the sources.

A multicomponent decomposition of the broad-band SED shows that relative contributions to the bolometric luminosity from the accretion disc, Comptonization and power-law components vary between sources with different luminosity and Hβ linewidth. We find a slight increase in contribution from the accretion disc as the luminosity increases, but a larger sample with more

sources at both low and high luminosities is needed to confirm this.

Our study also supports the distinctiveness of the NLS1s in the whole sample. We find that NLS1s tend to have a softer 2–10 keV spectrum, lower 2–10 keV luminosity, lower black hole mass, higher Eddington ratio and higher α_{ox} index. However, NLS1s do not stand out from the whole sample in terms of their bolometric luminosity distribution. We estimate the corona radius for every AGN in our sample from the SED fitting. This shows that on average NLS1s have smaller corona radii, and correspondingly a smaller coronal component contribution.

We compare the best-fitting black hole masses with those corrected for radiation pressure, and other estimates of black hole mass based on the R_{BLR}–L_{5100} relation, including numerous options for measuring the velocity width of the Hβ emission line. These results show that the black hole masses derived from SED fitting have a similar distribution to that derived from profiles corrected for radiation pressure effects, except for an offset of 0.3 dex lower in both the NLS1 and BLS1 subsamples. The black hole mass differences between NLS1s and BLS1s from these two methods (i.e. SED fitting and radiation pressure corrected profiles) are both smaller than inferred from other mass measurements. This implies that compared with black hole mass estimates based only on the Hβ FWHM, NLS1s may lie closer to the established M–σ_* relation at the low-mass end, when their black hole masses are corrected for radiation pressure, and when we use masses derived from our SED fitting.

Finally, we form three broad-band SED templates by co-adding SEDs in three subsamples (consisting of 12 objects each) to examine how the broad-band SED depends on Hβ FWHM velocity width, and by extending the Eddington ratio. The results show that there is a change in the SED shape as the FWHM increases, with NLS1s having the largest BBB in the extreme-UV region. Other important parameters such as $\Gamma_{2-10\,keV}$, $\kappa_{2-10\,keV}$ and α_{ox} also change as the Hβ FWHM increases. The implications of correlations between these parameters will be discussed in our next paper.

ACKNOWLEDGMENTS

CJ acknowledges financial support through the award of a Durham Doctoral Fellowship. This work is partially based on the data from the SDSS, whose funding is provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the US Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. This work is also partially based on observations obtained with *XMM–Newton*, an ESA science mission with instruments and contributions directly funded by ESA Member States and the USA (NASA). We also used *GALEX* data, which are based on observations made with the NASA *GALEX*. The *GALEX* is operated for NASA by the California Institute of Technology under NASA contract NAS5-98034.

REFERENCES

Antonucci R., 1993, ARA&A, 31, 473
Bentz M. C. et al., 2006, ApJ, 651, 775
Bentz M. C. et al., 2010, ApJ, 716, 993
Bessell M. S., 1991, A&A, 242, L17
Bian W., Huang K., 2010, MNRAS, 401, 507
Bian W., Zhao Y., 2004, MNRAS, 347, 607
Bian W., Hu C., Gu Q., Wang J., 2008, MNRAS, 390, 752
Boisson C., Joly M., Moultaka J., Pelat D., Serote Roos M., 2000, A&A, 357, 850
Boller T., Brandt W. N., Fink H., 1996, A&A, 305, 53
Boroson T. A., Green R. F., 1992, ApJS, 80, 109
Bouchet P., Lequeux J., Maurice E., Prevot L., Prevot-Burnichon M. L., 1985, A&A, 149, 330
Brocksopp C., Starling R. L. C., Schady P., Mason K. O., Romero-Colmenero E., Puchnarewicz E. M., 2006, MNRAS, 366, 953
Calzetti D., Armus L., Bohlin R. C., Kinney A. L., Koornneef J., Storchi-Bergmann T., 2000, ApJ, 533, 682
Canalizo G., Stockton A., 2001, ApJ, 555, 719
Canfield R. C., Puetter R. C., 1981, ApJ, 243, 390
Casebeer D. A., Leighly K. M., Baron E., 2006, ApJ, 637, 157
Collin S., Kawaguchi T., 2004, A&A, 426, 797
Collin S., Kawaguchi T., Peterson B. M., Vestergaard M., 2006, A&A, 456, 75
Collin-Souffrin S., Dumont A. M., 1990, A&A, 229, 292
Córdova F. A., Kartje J. F., Thompson R. J., Jr, Mason K. O., Puchnarewicz E. M., Harnden F. R., Jr, 1992, ApJS, 81, 661
Crenshaw D. M., Kraemer S. B., George I. M., 2003, ARA&A, 41, 117
Crummy J., Fabian A. C., Gallo L., Ross R. R., 2006, MNRAS, 365, 1067
Czerny B., Nikołajuk M., Różańska A, Dumont A.-M., Loska Z., Zycki P. T., 2003, A&A, 412, 317
Denney K. D. et al., 2010, ApJ, 721, 715
Done C., Gierliński M., 2005, Ap&SS, 300, 167
Done C., Kubota A., 2006, MNRAS, 371, 1216
Done C., Davis S., Jin C., Blaes O., Ward M., 2011, MNRAS, in press (doi:10.1111/j.1365-2966.2011.19779.x) (this issue)
Elvis M. et al., 1994, ApJS, 95, 1
Fitzpatrick E. L., 1986, AJ, 92, 1068
Gierliński M., Done C., 2004, MNRAS, 349, L7
Gierliński M., Zdziarski A. A., Poutanen J., Coppi P. S., Ebisawa K., Johnson W. N., 1999, MNRAS, 309, 496
Gierliński M., Middleton M., Ward M., Done C., 2008, Nat, 455, 369
Giveon U., Maoz D., Kaspi S., Netzer H., Smith P. S., 1999, MNRAS, 306, 637
Gladstone J., Roberts T., Done C., 2009, MNRAS, 397, 1836
Goodrich R. W., 1989, ApJ, 342, 224
Grandi S. A., 1982, ApJ, 255, 25
Grupe D., Mathur S., 2004, ApJ, 606, L41
Grupe D., Beuermann K., Thomas H. C., Mannheim K., Fink H. H., 1998, A&A, 330, 25
Grupe D., Beuermann K., Mannheim K., Thomas H. C., 1999, A&A, 350, 805
Haardt F., Maraschi L., 1991, ApJ, 380, L51
Hao H. et al., 2010, ApJ, 724, L59
Hu C., Wang J., Ho L. C., Chen Y., Bian W., Xue S., 2008, ApJ, 683, L115
Hubbard E. N., Puetter R. C., 1985, ApJ, 290, 394
Jin C., Done C., Ward M., Gierliński M., Mullaney J., 2009, MNRAS, 398, L16
Kalberla P. M. W., Burton W. B., Hartmann D., Arnal E. M., Bajaja E., Morras R., Pöppel W. G. L., 2005, A&A, 440, 775
Kaspi S., Smith P. S., Netzer H., Maoz D., Jannuzi B., Giveon U., 2000, ApJ, 533, 631
Kinney A. L., Calzetti D., Bohlin R. C., McQuade K., Storchi-Bergmann T., Schmitt H. R., 1996, ApJ, 467, 38
Komossa S., 2008, Rev. Mex. Astron. Astrofis. Ser. Conf., 32, 86
Kwan J., 1984, ApJ, 283, 70
Kwan J., Krolik J. H., 1981, ApJ, 250, 478
Lacy M., Sajina A., Petric A. O., Seymour N., Canalizo G., Ridgway S. E., Armus L., Storrie-Lombardi L. J., 2007, ApJ, 669, L61
Landt H., Elvis M., Ward M. J., Bentz M. C., Korista K. T., Karovska M., 2011, MNRAS, 414, 218
Lee J. C., Ogle P. M., Canizares C. R., Marshall H. L., Schulz N. S., Morales R., Fabian A. C., Iwasawa K., 2001, ApJ, 554, L13
Leighly K. M., 1999, ApJS, 125, 317
Lusso E. et al., 2010, A&A, 512, 34

Mannucci F., Basile F., Poggianti B. M., Cimatti A., Daddi E., Pozzetti L., Vanzi L., 2001, MNRAS, 326, 745
Maoz D. et al., 1993, ApJ, 404, 576
Marconi A., Axon D. J., Maiolino R., Nagao T., Pastorini G., Pietrini P., Robinson A., Torricelli G., 2008, ApJ, 678, 693
Mei L., Yuan W., Dong X., 2009, Res. Astron. Astrophys., 9, 269
Middleton M., Done C., 2010, MNRAS, 403, 9
Middleton M., Done C., Gierliński M., 2007, MNRAS, 381, 1426
Middleton M., Done C., Ward M., Gierliński M., Schurch N., 2009, MNRAS, 394, 250
Mihalas D., 1978, Stellar Atmospheres, 2nd edn. Freeman & Co., San Francisco, CA
Miller L., Turner T. J., Reeves J. N., 2008, A&A, 483, 437
Mullaney J. R., Ward M. J., Done C., Ferland G. J., Schurch N., 2009, MNRAS, 394, L16
Nandra K., Pounds K. A., 1994, MNRAS, 268, 405
Osterbrock D. E., 1989, Astrophysics of Gaseous Nebulae and Active Galacitc Nuclei. University Science Books, Mill Valley, CA
Osterbrock D. E., Pogge R. W., 1985, ApJ, 297, 166
Peterson B. M. et al., 2004, ApJ, 613, 682
Pigarov A. Y., Terry J. L., Lipschultz B., 1998, Plasma Phys. Control. Fusion, 40, 12
Prevot M. L., Lequeux J., Prevot L., Maurice E., Rocca-Volmerange B., 1984, A&A, 132, 389
Puchnarewicz E. M. et al., 1992, MNRAS, 256, 589
Puetter R. C., Levan P. D., 1982, ApJ, 260, 44
Seaton M. J., 1979, MNRAS, 187, 73
Shuder J. M., Osterbrock D. E., 1981, ApJ, 250, 55
Soria R., Puchnarewicz E. M., 2002, MNRAS, 329, 456
Sulentic J. W., Zwitter T., Marziani P., Dultzin-Hacyan D., 2000, ApJ, 536, L5
Talavera A., 2009, Ap&SS, 320, 177
Tsuzuki Y., Kawara K., Yoshii Y., Oyabu S., Tanabé T., Matsuoka Y., 2006, ApJ, 650, 57
Turner A. K., Fabian A. C., Lee J. C., Vaughan S., 2004, MNRAS, 353, 319
Turner A. K., Miller L., Reeves J. N., Kraemer S. B., 2007, A&A, 475, 121
Vasudevan R. V., Fabian A. C., 2007, MNRAS, 381, 1235 (VF07)
Vasudevan R. V., Fabian A. C., 2009, MNRAS, 392, 1124
Verner E., Bruhweiler F., Johansson S., Peterson B., 2009, Phys. Scr., 2009, T134
Véron-Cetty M. P., Joly M., Véron P., 2004, A&A, 417, 515
Walter R., Fink H. H., 1993, A&A, 274, 105
Wang T., Lu Y., 2001, A&A, 377, 52
Ward M., Elvis M., Fabbiano G., Carleton N. P., Willner S. P., Lawrence A., 1987, ApJ, 315, 74
Wills B. J., Netzer H., Wills D., 1985, ApJ, 288, 94
Woo J., Urry C. M., 2002, ApJ, 579, 530
Zdziarski A. A., Poutanen J., Johnson W. N., 2000, ApJ, 542, 703
Zhou H., Wang T., Yuan W., Lu H., Dong X., Wang J., Lu J., 2006, ApJS, 166, 128
Zhu L., Zhang S., Tang S., 2009, ApJ, 700, 1173

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure A1. The spectral fitting results.
Table B1. *XMM–Newton* and SDSS DR7 source positions and separations of our sample.
Table C1. Black hole masses from different methods.
Table D1. Emission-line parameters for the whole sample.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

This paper has been typeset from a TEX/LATEX file prepared by the author.